

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C,
xy

DR WPI: 2001-041424/05.
 DR N-PSDB; AAF27729.
 PT Isolated polypeptide with a human transport protein sequence is useful
 PT for the diagnosis, prevention and treatment of disorders associated
 PT with the immune, reproductive and cardiovascular systems -
 XX
 PS Claim 2, Page 130, 165pp; English.
 XX
 CC The present invention provides the protein and coding sequences for 43
 CC novel human transport proteins (designated TPPTS). These can be used in
 CC the diagnosis and treatment of transport, metabolic, neurological,
 CC reproductive, cardiovascular and immune disorders, and cell proliferative
 CC disorders such as cancer.
 CC
 XX
 SQ Sequence 219 AA;
 Query Match 100.0%; Score 219; DB 22; Length 219;
 Best Local Similarity 100.0%; Pred. No. 1,1e-200;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKLLMACIVCAVFAKRRPFPIGEDNDGHLHSLNIPYGINLPPLYYRPVNTVP 60
 DB 1 MKLLMACIVCAVFAKRRPFPIGEDNDGHLHSLNIPYGINLPPLYYRPVNTVP 60
 QY 61 SYPGNTYDTGCLPSYFWILTSRGFPYVYHIRGFPPLATQLVNPLPRGFPVPPSRFSA 120
 DB 61 SYPGNTYDTGCLPSYFWILTSRGFPYVYHIRGFPPLATQLVNPLPRGFPVPPSRFSA 120
 QY 121 AAPPAPPIAEPAPAAAPLTATPVAAEPAPAGAVAAEPAAEPVGAEPAAEPVAAEPAA 180
 DB 121 AAPPAPPIAEPAPAAAPLTATPVAAEPAPAGAVAAEPAAEPVGAEPAAEPVAAEPAA 180
 QY 181 EAPVGEPAABEPSPAEPATAKPADEPHSPSLAQANQ 219
 DB 181 EAPVGEPAABEPSPAEPATAKPADEPHSPSLAQANQ 219
 DB
 RESULT 2
 AAE27862
 ID AAE27862 standard; Protein; 219 AA.
 XX
 AC AAE27862;
 XX
 DT 27-DEC-2002 (first entry)
 XX
 DE Human zsig63 protein.
 XX
 KW Human; secreted salivary protein; zsig63 protein; host defense protein;
 KW immune modulating factor; antipathogenic; cell-cell signalling molecule;
 KW growth factor; cytokine; growth factor hormone activity; dental caries;
 KW infection; tooth decay; periodontal disease; gastrointestinal disease;
 KW thrush; urinary tract infection; vaginal infection; diabetes; obesity;
 KW anti-inflammatory; chronic tissue damage; lung dysfunction; restenosis;
 KW gene therapy; salivary gland dysfunction; prostate gland dysfunction;
 KW forensic DNA profiling; chondrosarcoma; atherosclerosis; chromosome 4.
 XX
 OS Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Reptide 1..15
 FT /label= Signal_peptide
 FT Protein 16..219
 FT /note= "Mature human zsig63 protein"
 FT Region 14..21
 FT /note= "Hydrophilic region"
 FT Domain 16..37
 FT /note= "Domain 1"
 FT Region 17..33
 FT /note= "Antigenic epitope"
 FT Region 24..30
 FT /note= "Hydrophilic region"
 FT Domain 38..126

FT /note= "Domain 2"
 FT Region 66..73
 FT /note= "Antigenic epitope"
 FT Region 103..108
 FT /note= "Antigenic epitope"
 FT Region 124..133
 FT /note= "Repeat 1"
 FT Domain 127..219
 FT /note= "Domain 3"
 FT Region 134..138
 FT /note= "Repeat 2"
 FT Region 139..143
 FT /note= "Repeat 3"
 FT Region 144..148
 FT /note= "Repeat 4"
 FT Region 149..153
 FT /note= "Repeat 5"
 FT Region 154..158
 FT /note= "Repeat 6"
 FT Region 159..163
 FT /note= "Repeat 7"
 FT Region 164..168
 FT /note= "Repeat 8"
 FT Region 169..173
 FT /note= "Repeat 9"
 FT Region 174..178
 FT /note= "Repeat 10"
 FT Region 179..183
 FT /note= "Repeat 11"
 FT Region 184..188
 FT /note= "Repeat 12"
 FT Region 187..192
 FT /note= "Hydrophilic region"
 FT Region 189..193
 FT /note= "Repeat 13"
 FT Region 190..197
 FT /note= "Antigenic epitope"
 FT Region 194..198
 FT /note= "Repeat 14"
 FT Region 199..203
 FT /note= "Repeat 15"
 FT Region 202..215
 FT /note= "Antigenic epitope"
 FT Region 204..208
 FT /note= "Repeat 16"
 XX
 PN US2002090677-A1.
 XX
 PD 11-JUL-2002.
 XX
 PF 03-AUG-2001; 2001US-0923236.
 XX
 XX 17-MAR-1999; 99US-124820P.
 PR 17-MAR-2000; 2000US-0527345.
 XX
 PA (ADLER/) ADLER D A.
 PA (SHEP/) SHEPPARD P O.
 XX
 PI Adler DA, Sheppard PO;
 XX
 DR WPI: 2002-642378/69.
 DR N-PSDB; AAD45050.
 PT Novel secreted salivary polypeptide, zsig63, useful as antimicrobial
 PT agent for treating microbial infection, dental caries, periodontal
 PT disease, thrush gastrointestinal disease, and for aiding digestion -
 XX
 PS Claim 10; Page 28; 33pp; English.
 XX
 CC The invention relates to human secreted salivary polypeptide designated
 CC as zsig63 and nucleic acid molecules encoding such polypeptides. zsig63
 CC can be used in detecting agonists and antagonists of its activity, and
 CC is also useful as a host defense polypeptide, immune modulating factor,

CC antipathogenic polypeptide, cell-cell signalling molecule, growth factor,
CC cytokine, or as secreted extracellular matrix associated proteins with
CC growth factor hormone activity. It is useful for treating conditions
CC associated with pathological microbes, including bacterial, fungal and
CC viral infections, for treating dental caries (tooth decay), periodontal
CC disease, thrush and gastrointestinal disease, for treating urinary tract
CC infection, vaginal infection and for preventing infection in skin and
CC other epithelial wounds. zsig63 is useful for establishing normal
CC microflora and protect against pathogenic colonisation and invasion, for
CC treating chronic tissue damage e.g. damage in extremities associated
CC with diabetes and useful as anti-inflammatory agents. It is useful as a
CC marker of lung dysfunction, salivary gland dysfunction, or dysfunction of
CC prostate gland. It is also therapeutically useful for aiding digestion.
CC Polynucleotides of the invention are used in gene therapy for increasing
CC or inhibiting zsig63 activity, for detecting abnormalities on human
CC chromosome 4 associated with disease or other human traits and as
CC diagnostic in forensic DNA profiling. Sequences of the invention are
CC useful for stimulating proliferation or differentiation of cardiac
CC myocytes, for proliferation or differentiation of adipocytes and for
CC inhibiting chondrosarcomas, atherosclerosis, restenosis and obesity.
CC The present sequence is human zsig63 protein. zsig63 gene is located
CC on chromosome 4.
XX
XX
SQ Sequence 219 AA:

Query Match 100.0%; Score 219; DB 23; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.1e-200;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLLIMACTVCAVFAARKRFPFEGEDDNDGHLPLPSLNIPIYGINLPEPLYYRPVNTVP 60
Db 1 MKLLIMACTVCAVFAARKRFPFEGEDDNDGHLPLPSLNIPIYGINLPEPLYYRPVNTVP 60
QY 61 SYPGNTYTDGTPSPYFWITSGPPYVYHIRGPIPLATQNLNPPDPKGFPPVPSRFPSSA 120
Db 61 SYPGNTYTDGTPSPYFWITSGPPYVYHIRGPIPLATQNLNPPDPKGFPPVPSRFPSSA 120
QY 121 AAAPAPPIAAEPAAAPITATPVAAEPAGAVAAEPAAEPVGAEPAAEPVAAEPAA 180
Db 121 AAAPAPPIAAEPAAAPITATPVAAEPAGAVAAEPAAEPVGAEPAAEPVAAEPAA 180
QY 181 EAPVGEPAAEPSPAEPATAKPAEPHPSPSLQANQ 219
Db 181 EAPVGEPAAEPSPAEPATAKPAEPHPSPSLQANQ 219

RESULT 3
ABG31608
ID ABG31608 standard; Protein; 219 AA.
XX
XX ABG31608;
DT 15-NOV-2002 (first entry)
XX
DE Human secreted salivary protein zsig63.
XX
KM Human, secreted salivary protein; zsig63, immunogen; zsig63-cytokine;
KM antibody-cytokine; in vivo killing; pathological microbe; bacteria;
KM fungal; viral; infection; salivary gland; anti-microbial; dental caries;
KM tooth decay; periodontal disease; thrush; gastrointestinal disease;
KM urinary tract infection; vaginal infection; skin infection; microflora;
KM epithelial wound; pathogenic colonisation; invasion; pro-inflammatory;
KM chronic tissue damage; vascular system; diabetes; anti-inflammatory;
KM incompetent immune system; AIDS; acquired immunodeficiency syndrome;
KM chemotherapy; radiation treatment; lung infection; cystic fibrosis;
KM digestion; chromosome 4.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Region 1..15
FT /note= "Fusion protein peptide, specifically claimed
FT in claim 18"
FT

FT Region 1..219
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 9..204
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 14..19
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 16..21
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 16..37
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 16..219
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 17..33
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 24..29
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 24..33
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 25..30
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 38..126
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 66..73
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 103..108
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 109..215
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 124..133
FT /label= Repeat_1
FT 127..219
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT Region 134..138
FT /label= Repeat_2
FT 139..143
FT /label= Repeat_3
FT 144..148
FT /label= Repeat_4
FT 149..153
FT /label= Repeat_5
FT 154..158
FT /label= Repeat_6
FT 159..163
FT /label= Repeat_7
FT 164..168
FT /label= Repeat_8
FT 169..173
FT /label= Repeat_9
FT 174..178
FT /label= Repeat_10
FT 179..183
FT /label= Repeat_11
FT 184..188
FT /label= Repeat_12
FT 187..192
FT /note= "Antigenic peptide, specifically claimed in
FT claim 15"
FT 189..193
FT Region

FT FT /label= Repeat_13
 FT Region 190..197
 FT /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 194..198
 FT /label= Repeat_14
 FT Region 199..203
 FT /label= Repeat_15
 FT 202..215
 FT /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 204..208
 FT /label= Repeat_16
 FT US2002061701-A1.
 PD 27-JUN-2002.
 XX
 PD 03-AUG-2001; 2001US-0922480.
 PF 17-MAR-1999; 99US-124820P.
 PR 17-MAR-2000; 2000US-0527345.
 XX
 PA (ADLER/) ADLER D A.
 XX (SHEPPARD/) SHEPPARD P O.
 PI Adler DA, Sheppard PO;
 XX WPI; 2002-635468/68.
 DR N-PSDB; ABS52633, ABS52634.
 XX
 PT Novel secreted salivary protein, zsig63 and polynucleotide encoding it
 PT useful for treating microbial infections, inflammatory conditions,
 PT dental caries and lung infections associated with cystic fibrosis
 XX
 PS Claim 10; Page 28; 33pp; English.

CC The present invention relates to a new secreted salivary protein, zsig63.
 CC The invention is useful for detecting in a test sample, the presence of
 CC an antagonist or agonist of zsig63 protein activity. The invention is
 CC also useful as an immunogen for producing an antibody to zsig63
 CC polypeptide, zsig63-cytokine fusion proteins or antibody-cytokine fusion
 CC protein are useful for enhancing in vivo killing of target tissues.
 CC Pharmaceutical composition comprising purified zsig63 polypeptide are
 CC useful in the treatment of conditions associated with pathological
 CC microbes, including bacterial, fungal and viral infections. High
 CC expression of zsig63 in salivary gland suggests that anti-microbial
 CC polypeptides are useful for treatment of dental caries (tooth decay),
 CC periodontal disease, thrush and gastrointestinal disease. Other
 CC applications can be used in urinary tract infections, vaginal infections,
 CC prevention of infection in skin and other epithelial wounds. The
 CC polypeptides can be used to establish normal microflora and protect
 CC against pathogenic colonisation and invasion. The invention is useful
 CC when pro-inflammatory activity is desired. Applications for
 CC such pro-inflammatory activity include the treatment of chronic tissue
 CC damage, particularly in areas having a limited or damaged vascular system
 CC e.g., damage in extremities associated with diabetes. Antagonists to
 CC zsig63 polypeptides may be useful as anti-inflammatory agents. The
 CC invention is useful for the treatment of patients having incompetent
 CC immune system, such as AIDS (acquired immunodeficiency syndrome) patients
 CC or individuals that have undergone chemotherapy, radiation treatment. The
 CC invention is also useful for the treatment of lung infections associated
 CC with cystic fibrosis and its agonists or antagonists are useful for
 CC aiding digestion. The present amino acid sequence represents the human
 CC secreted salivary protein zsig63 of the invention. This sequence is
 CC encoded by the human zsig63 gene located on chromosome 4.
 XX
 SQ Sequence 219 AA;

Query Match 100.0%; Score 219; DB 23; Length 219;
 Best Local Similarity 100.0%; Pred. No. 1.1e-200;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLLIMACTVCVAFARKRPPPIGEDNDGDGPHLPSINIPYGIKMLPPLYYREVENTVP 60
 Db 1 MKLLIMACTVCVAFARKRPPPIGEDNDGDGPHLPSINIPYGIKMLPPLYYREVENTVP 60
 QY 61 SYPGNTYTDGTPSYWMLTSPGPPYVYHNGFPLATOLNPLPDRGPPFPSPRSFSA 120
 Db 61 SYPGNTYTDGTPSYWMLTSPGPPYVYHNGFPLATOLNPLPDRGPPFPSPRSFSA 120
 QY 121 AAAPAAPPIAAEPAAAALPLTTPVAAEPAGAPVAAEPAAEPVGAEPAAEPVAAEPAA 180
 Db 121 AAAPAAPPIAAEPAAAALPLTTPVAAEPAGAPVAAEPAAEPVGAEPAAEPVAAEPAA 180
 QY 181 EAPVGEPAAEPSPAEPATAKPAAPBPSPSLAQANQ 219
 Db 181 EAPVGEPAAEPSPAEPATAKPAAPBPSPSLAQANQ 219

RESULT 4
 AAU74536
 ID AAU74536 standard; Protein; 219 AA.
 XX
 AC AAU74536;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Human zsig63 polypeptide.
 XX
 KW Human; zsig63; chromosome 4q12-q13; salivary protein; antimicrobial;
 KW microbial infection; tooth decay; periodontal disease; thrush; emphysema;
 KW gastrointestinal disease; urinary tract infection; vaginal infection;
 KW skin infection; epithelial wound; chronic tissue damage; cystic fibrosis;
 KW acquired immunodeficiency syndrome; AIDS; lung infection; sarcoidosis;
 KW chronic bronchitis; gene therapy; protein therapy.
 XX
 OS Homo sapiens.
 XX
 PN US6331413-B1.
 XX
 PD 18-DEC-2001.
 XX
 PF 17-MAR-2000; 2000US-0527345.
 XX
 PR 17-MAR-1999; 99US-124820P.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Adler DA, Sheppard PO;
 XX
 DR WPI; 2002-096707/13.
 XX
 DR N-PSDB; AAS20591.
 XX
 PT Polynucleotides encoding salivary proteins useful as anti-microbial
 PT agents -
 XX
 PS Claim 1; Column 49-52; 29pp; English.

CC The invention relates to a polynucleotide derived from the 4q12-4q13
 CC region of human chromosome 4 and encoding a zsig63 polypeptide, a
 CC secreted salivary protein with anti-microbial activity. Due to their
 CC microbial activity, the sequences can be used in the study of microbial
 CC infections, e.g. for recombinant production of anti-microbial proteins.
 CC The sequences can be used in the treatment of tooth decay, periodontal
 CC disease, thrush, gastrointestinal disease, urinary tract infections,
 CC vaginal infections, skin infections, epithelial wounds, chronic tissue
 CC damage, acquired immunodeficiency syndrome (AIDS), cystic fibrosis, lung
 CC infections, sarcoidosis, emphysema and chronic bronchitis. This sequence
 CC represents human zsig63.
 XX
 SQ Sequence 219 AA;

Query Match 100.0%; Score 219; DB 23; Length 219;
 Best Local Similarity 100.0%; Pred. No. 1.1e-200;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MKLIMACIVCAFAKRRRPFPIGEDNDNDGHLPSLNPYGIKRLPPPLYRPPVNTVP 60
Db      1 MKLIMACIVCAFAKRRRPFPIGEDNDNDGHLPSLNPYGIKRLPPPLYRPPVNTVP 60
Qy      61 SYPGNTYDTGGLSPYMWILTSPPFPYVHINRFPPLATQINVPPLPPRGFPFPPSPRFTSA 120
Db      61 SYPGNTYDTGGLSPYMWILTSPPFPYVHINRFPPLATQINVPPLPPRGFPFPPSPRFTSA 120
Qy      121 AAAPAPPIAAEPAAAPLTATPVAAPPAAGPVAAPPAEAPVGAEPAAEPVAAEPAA 180
Db      121 AAAPAPPIAAEPAAAPLTATPVAAPPAAGPVAAPPAEAPVGAEPAAEPVAAEPAA 180
Qy      181 EAPVGEPAABEPSPAPPAETAKAPPEHPSPSLEQANQ 219
Db      181 EAPVGEPAABEPSPAPPAETAKAPPEHPSPSLEQANQ 219

RESULT 5
ABU08515
ID      ABU08515 standard; Protein; 219 AA.
XX      ABU08515;
AC      28-MAY-2003 (first entry)
XX      Human zsig63 protein.
DE      Human; immunogen; zsig63; adhesion; salivary gland; dental carries;
XX      periodontal disease; thrush; gastrointestinal disease; epithelial wound;
XX      urinary tract infection; vaginal infection; skin infection;
XX      pro-inflammatory; chronic tissue damage; vascular system; diabetes; AIDS;
XX      lung infection; cystic fibrosis; lung dysfunction; digestive;
XX      salivary gland carcinoma; Pneumocystis carinii infection; emphysema;
XX      chronic bronchitis; prostate dysfunction; prostate adenocarcinoma;
XX      cell culture media; gene therapy; human chromosome 4q12-4q13;
XX      dentinogenesis imperfecta; dentin dysplasia type II.
OS      Homo sapiens.
XX      Key
XX      Location/Qualifiers
XX      1..15
XX      /label= Signal peptide
XX      /note= "Specifically claimed in claim 8"
XX      16..219
XX      /label= Mature_zsig63
XX      /note= "Specifically claimed in claim 1"
XX      16..37
XX      /label= Domain_1
XX      /note= "Specifically claimed in claim 1"
XX      38..126
XX      /label= Domain_2
XX      /note= "Specifically claimed in claim 1"
XX      127..219
XX      /label= Domain_3
XX      /note= "Specifically claimed in claim 1"
XX      14..19
XX      /label= Immunogenic_peptide
XX      /note= "Specifically claimed in claim 15"
XX      14..19
XX      /label= Immunogenic_peptide
XX      /note= "Specifically claimed in claim 15"
XX      16..21
XX      /label= Immunogenic_peptide
XX      /note= "Specifically claimed in claim 15"
XX      24..29
XX      /label= Immunogenic_peptide
XX      /note= "Specifically claimed in claim 15"
XX      25..30
XX      /label= Immunogenic_peptide
XX      187..192
XX      /label= Immunogenic_peptide
XX      /note= "Specifically claimed in claim 15"
XX      24..33
XX      /label= Immunogenic_peptide
XX      /note= "Specifically claimed in claim 15"
XX      17..33
XX      /label= Immunogenic_peptide
XX      /note= "Specifically claimed in claim 15"
XX      66..73
XX      /label= Immunogenic_peptide
XX      /note= "Specifically claimed in claim 15"
XX      103..108
XX      /label= Immunogenic_peptide
XX      /note= "Specifically claimed in claim 15"
XX      190..197
XX      /label= Immunogenic_peptide
XX      /note= "Specifically claimed in claim 15"
XX      202..215
XX      /label= Immunogenic_peptide
XX      /note= "Specifically claimed in claim 15"
XX      190..215
XX      /label= Immunogenic_peptide
XX      /note= "Specifically claimed in claim 15"
XX      124..133
XX      /label= Repeat_1
XX      134..138
XX      /label= Repeat_2
XX      139..143
XX      /label= Repeat_3
XX      144..148
XX      /label= Repeat_4
XX      149..153
XX      /label= Repeat_5
XX      154..158
XX      /label= Repeat_6
XX      159..163
XX      /label= Repeat_7
XX      164..168
XX      /label= Repeat_8
XX      169..173
XX      /label= Repeat_9
XX      174..178
XX      /label= Repeat_10
XX      179..183
XX      /label= Repeat_11
XX      184..188
XX      /label= Repeat_12
XX      189..193
XX      /label= Repeat_13
XX      194..198
XX      /label= Repeat_14
XX      199..203
XX      /label= Repeat_15
XX      204..208
XX      /label= Repeat_16
XX      US2002173027-A1.
XX      21-NOV-2002.
XX      03-AUG-2001; 2001US-0922469.
XX      17-MAR-1999; 99US-124820P.
XX      17-MAR-2000; 2000US-0527345.
XX      (ADLER/) ADLER D A.
XX      (SHEP/) SHEPPARD P O.
XX      Adler DA, Sheppard PO;
XX      WPI; 2003-326428/31.

```

DR N-PSDB; ABX93594.

XX Novel isolated zsig63 polypeptide, member of the adhesin family, useful
PT for treating dental caries, periodontal disease, thrush,
PT gastrointestinal disease, urinary tract infections, vaginal infections,
PT skin infections

PS Claim 10; Page 27-28; 32pp; English.

XX The invention relates to an isolated zsig63 polypeptide comprising at
XX least 90% identity to an amino acid sequence which comprises domain 1 of
XX zsig63, domain 2, domain 3, mature zsig63 and full length zsig63. Also
XX included are the polynucleotide encoding zsig63, a zsig63 expression
XX vector, a cultured cell comprising the vector and expressing the protein,
XX a DNA encoding a fusion protein (comprising amino acids 1-15, 16-37,
XX 38-126, 127-219 or 16-219 of zsig63 and an additional protein), using a
XX zsig63 reporter gene construct to identify zsig63 agonists, and
XX producing an anti-zsig63 antibody using zsig63 immunogenic peptides,
XX zsig63 is useful for detecting in a test sample, the presence of
XX antagonist of zsig63 protein activity. Zsig63 has antimicrobial
XX activity and since exhibits high expression in salivary gland, can be
XX used for treating dental caries, periodontal disease, thrush, and
XX gastrointestinal disease, urinary tract infections, vaginal infections,
XX skin infections and other epithelial wounds. The polypeptides can be
XX used to establish normal microflora and protect against pathogenic
XX colonization and invasion. Zsig63 can also be used for providing
XX pro-inflammatory activity for treating chronic, tissue damage
XX particularly in areas having limited or damaged vascular system, e.g.
XX in diabetes, and for treating immunocompromised AIDS patients or in
XX individuals that have undergone chemotherapy, radiation treatment, for
XX treating lung infections e.g. in cystic fibrosis. Detection of zsig63
XX polypeptide at relatively high levels in the trachea may indicate that
XX such polypeptides may serve as a marker of lung dysfunction. Zsig63 is
XX also useful in diagnosing conditions associated with salivary gland or
XX lung dysfunction including salivary gland carcinoma, Pneumocystis carinii
XX infection, emphysema, chronic bronchitis, prostate dysfunctions such
XX as prostate adenocarcinoma, aiding digestion, and as components of
XX defined cell culture media and may be used to replace serum that is
XX commonly used in culture. The DNA is useful in gene therapy applications
XX to increase or inhibit zsig63 activity, and for detecting abnormalities
XX on human chromosome 4 (e.g. 4q12-4q13, associated with dentinogenesis
XX imperfecta, and dentin dysplasia type II). Zsig63 is an adhesin family
XX member. The present sequence represents human zsig63.

XX Sequence 219 AA;

XX Query Match 100.0%; Score 219; DB 24; Length 219;

XX Best Local Similarity 100.0%; Pred. No. 1.1e-200;

XX Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLLMAGTVCVAFKRRPPFPGEDNDGPHLSNIPYIRNLPPLYRRPNVTP 60
DB 1 MKLLMAGTVCVAFKRRPPFPGEDNDGPHLSNIPYIRNLPPLYRRPNVTP 60
QY 61 STPGNTYDTGTPSYWIMLTSPGFVYVHIGRPLATOLNVPPLPRGPFVPPSRFFSA 120
DB 61 STPGNTYDTGTPSYWIMLTSPGFVYVHIGRPLATOLNVPPLPRGPFVPPSRFFSA 120
QY 121 AAAAAPIAEEPAAPLTAFTVAEPAAGVAAEPAEPAEPAEPAEPAEPAEPA 180
DB 121 AAAAAPIAEEPAAPLTAFTVAEPAAGVAAEPAEPAEPAEPAEPAEPAEPA 180
QY 181 EAPVGEPAEPEEPAEPAEPAEPAEPAEPAEPAEPAEPAEPAEPAEPAEPA 219
DB 181 EAPVGEPAEPEEPAEPAEPAEPAEPAEPAEPAEPAEPAEPAEPAEPAEPA 219

RESULT 6
AA19472
ID AA19472 standard; Protein; 221 AA.
XX
AC AA19472;
XX

DT 14-JUL-1999 (first entry)

XX Amino acid sequence of a human secreted protein.

XX Human secreted protein; cancer; tumour; neurodegenerative disorder;
XX developmental abnormality; fetal deficiency; blood disorder; leukemia;
XX immune system disease; autoimmune disease; hepatic disease;
XX renal disease; lymphoma; inflammation; allergy; ischemic shock;
XX Alzheimer's; cognitive disorder; schizophrenia; prostate disease;
XX obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease;
XX lung disease; thymus disease; digestive disorder; endocrine disorder;
XX infection; AIDS.

XX Homo sapiens.

XX WO922243-A1.

XX 06-MAY-1999.

XX 23-OCT-1998; 98WO-US22376.

XX 24-OCT-1997; 97US-0063387.

XX 24-OCT-1997; 97US-0062784.

XX 24-OCT-1997; 97US-0063088.

XX 24-OCT-1997; 97US-0063089.

XX 24-OCT-1997; 97US-0063090.

XX 24-OCT-1997; 97US-0063091.

XX 24-OCT-1997; 97US-0063092.

XX 24-OCT-1997; 97US-0063097.

XX 24-OCT-1997; 97US-0063098.

XX 24-OCT-1997; 97US-0063100.

XX 24-OCT-1997; 97US-0063101.

XX 24-OCT-1997; 97US-0063109.

XX 24-OCT-1997; 97US-0063110.

XX 24-OCT-1997; 97US-0063111.

XX 24-OCT-1997; 97US-0063148.

XX 24-OCT-1997; 97US-0063386.

XX (HUMA-) HUMAN GENOME SCT INC.

XX Brewer LA, Carter KC, Duan DR, Ebner R, Endress GA;

XX Feng P, Florence C, Florence KA, Greene JM, Janat F;

XX Kay H, Lafleur DM, Moore PA, Ni J, Olsen HS, Rosen CA;

XX Ruben SM, Shi Y, Soppet DR, Wei Y, Young P;

XX WPI; 1999-303069/25.

XX N-PSDB; AA19472.

XX New isolated human genes and the secreted polypeptides they encode

XX Claim 11; Page 401-402; 546pp; English.

XX The specification describes cDNA sequences (AA19472-19470) encoding
XX human secreted proteins (AA19472-19470). The polynucleotides and their
XX corresponding secreted polypeptides are useful for preventing, treating
XX or ameliorating medical conditions, e.g. by protein or gene therapy.
XX Pathological conditions can also be diagnosed by determining the amount
XX of the polypeptides in a sample or by determining the presence of
XX mutations in the polynucleotides. Specific uses are described for each
XX of the polynucleotides, based on which tissues they are most highly
XX expressed in, and include developing products for the diagnosis or
XX treatment of cancer, tumours, neurodegenerative disorders, developmental
XX abnormalities and fetal deficiencies, blood disorders, leukemia,
XX diseases of the immune system, autoimmune diseases, hepatic and renal
XX disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's
XX and cognitive disorders, schizophrenia, prostate diseases, obesity,
XX disorders involving osteoclasts such as osteoporosis, arthritis or
XX malignancies, diseases of testes, lung or thymus, digestive/endocrine
XX disorders, infections and AIDS. The polypeptides are also useful for
XX identifying their binding partners.

XX Sequence 221 AA;

Query Match 68.5%; Score 150; DB 20; Length 221;
Best Local Similarity 100.0%; Pred. No. 7.4e-135;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLLMACIVCAVAPARKRRPFPIGEDDDNDGHPHPSLNIPYGINRLPPELYRPNVTVP 60
DB 1 MKLLMACIVCAVAPARKRRPFPIGEDDDNDGHPHPSLNIPYGINRLPPELYRPNVTVP 60
QY 61 SYPGNTYTDGTPSYPMWLTSGPFYVYHIRGFPPLATQLVNPPLPGRGFPVPSRPFSA 120
DB 61 SYPGNTYTDGTPSYPMWLTSGPFYVYHIRGFPPLATQLVNPPLPGRGFPVPSRPFSA 120
QY 121 AAPAPPIAABPAAAPLTATPVAAEPAA 150
DB 121 AAPAPPIAABPAAAPLTATPVAAEPAA 150

RESULT 7
AAM30653
ID AAM30653 standard; Protein; 219 AA.
AC AAM30653;
DT 12-APR-1999 (first entry)
DE Human secreted protein clone cp16 1 protein.
KW Human; secreted protein; nutritional activity; cytokine; vaccine;
KW cell proliferation; differentiation; immune stimulation; suppression;
KW haematopoiesis regulation; tissue growth; activin; inhibitor; chemotactic;
KW chemokinetic; haemostatic; thrombolytic; anti-inflammatory; gene therapy;
KW tumour invasion suppression; tumour inhibition.
OS Homo sapiens.
XX MO9901466-A1.
XX 14-JAN-1999.
XX 01-JUL-1998; 98WO-US13813.
XX 27-OCT-1997; 97US-0958304.
XX 02-JUL-1997; 97US-0887195.
XX (GENY) GENETICS INST INC.
XX PA
XX AGostino MJ, Jacobs K, Lavallie ER, McCoy JM, Racie LA;
PI Spaulding V, Treacy M;
XX WPI; 1999-105994/09.
XX N-PSDB; AAV80740.
XX
XX New polynucleotides encoding secreted human proteins - are derived
XX from human foetal brain, adult testes, adult brain, foetal kidney,
XX adult salivary gland, or adult blood cDNA libraries, useful as, e.g.
XX potential vaccines
XX
XX Claim 24; Page 71-72; 107pp; English.
XX
XX The present sequence represents a human secreted protein from clone
XX cp16 1, deposited as ATCC 98482. Human secreted protein clone
XX polynucleotides and proteins are predicted to have biological
XX activities which would make them suitable for treating, preventing or
XX ameliorating medical conditions in humans and animals. Suggested
XX activities include nutritional activity, cytokine and cell
XX proliferation/differentiation activity, immune stimulating (e.g. as
XX vaccine) or suppressing activity, haematopoiesis regulating activity,
XX tissue growth activity, activin/inhibin activity, chemotactic/
XX chemokinetic activity, haemostatic and thrombolytic activity, receptor/
XX ligand activity, anti-inflammatory activity, cadherin/tumour invasion
XX suppressor activity, and tumour inhibition activity. The polynucleotides
XX are also stated to be useful for gene therapy.

XX SQ Sequence 219 AA;
Query Match 63.9%; Score 140; DB 20; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.5e-125;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLLMACIVCAVAPARKRRPFPIGEDDDNDGHPHPSLNIPYGINRLPPELYRPNVTVP 60
DB 1 MKLLMACIVCAVAPARKRRPFPIGEDDDNDGHPHPSLNIPYGINRLPPELYRPNVTVP 60
QY 61 SYPGNTYTDGTPSYPMWLTSGPFYVYHIRGFPPLATQLVNPPLPGRGFPVPSRPFSA 120
DB 61 SYPGNTYTDGTPSYPMWLTSGPFYVYHIRGFPPLATQLVNPPLPGRGFPVPSRPFSA 120
QY 121 AAPAPPIAABPAAAPLT 140
DB 121 AAPAPPIAABPAAAPLT 140

RESULT 8
ABB64202
ID ABB64202 standard; Protein; 75 AA.
AC ABB64202;
DT 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 19398.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX N-PSDB; ABL08305.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosure: SEQ ID NO 19398; 21pp + Sequence listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (AB857737-AB872072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 75 AA;
Query Match 4.1%; Score 9; DB 22; Length 75;
Best Local Similarity 100.0%; Pred. No. 0.91;

273 0933

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAAP 128
 |||||
 45 AAAAPAAP 53

RESULT 9
 ID AAM03688 standard; Protein; 322 AA.
 AC AAM03688;
 DT 09-MAR-1997 (first entry)
 DE Leishmania chagasi acidic ribosomal antigen LcPO.
 KW Leishmania chagasi; acidic ribosomal antigen; LcPO;
 XX epitope; K39.
 OS Leishmania chagasi.
 XX
 XX Key Location/Qualifiers
 FT CDS 30..1202
 FT /*tag= a
 XX
 PN MO9633414-A2.
 XX
 XX 24-OCT-1996.
 PD
 XX 19-APR-1996; 96WO-US05472.
 PF
 XX 21-APR-1995; 95US-0428414.
 PR
 XX (CORI-) CORIXA CORP.
 PA
 XX Reed SG;
 PI
 XX WPI; 1996-485884/48.
 DR
 XX N-PSDB; AAT42164.
 DR
 XX New Leishmania acidic ribosomal P-protein family polypeptide - used
 PT to develop prods. for diagnosis, detection and protection against
 PT Leishmania infections
 XX
 XX Disclosure; Page 29-32; 76pp; English.
 PS
 XX
 XX Compounds including polypeptides that contain at least an epitope of
 CC the L. chagasi acidic ribosomal antigen LcPO are useful in a variety
 CC of immunoassays for detecting Leishmania infection. Portions of
 CC LcPO (AAT42164) consg. at least the 17 C-terminal amino acids (AAT42165)
 CC have been found to generate a signal in an ELISA that is equivalent
 CC to that generated by the full length LcPO. A combination
 CC polypeptide may also be used, comprising an LcPO epitope along with
 CC an epitope derived from the Leishmania K39 antigen (AAT42166), pref.
 CC the K39 repeat unit antigen having the sequence given in AAM03689.
 XX
 XX Sequence 322 AA;
 SQ

Query Match 4.1%; Score 9; DB 17; Length 322;
 Best Local Similarity 100.0%; Pred. No. 3.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 AAEPAAP 138
 |||||
 289 AAEPAAP 297

RESULT 10
 ID ABB65430 standard; Protein; 728 AA.
 AC ABB65430;
 XX

XX 26-MAR-2002 (first entry)
 DT
 XX
 XX Drosophila melanogaster polypeptide SEQ ID NO 23082.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR
 XX 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI; 2001-656860/75.
 DR
 XX N-PSDB; ABL09533.
 DR
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 XX
 PS Disclosure; SEQ ID NO 23082; 21pp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 728 AA;
 SQ

Query Match 4.1%; Score 9; DB 22; Length 728;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 119 SAAAPAAP 127
 |||||
 418 SAAAPAAP 426

RESULT 11
 ID AAB76532 standard; Protein; 1874 AA.
 XX
 XX AAB76532;
 AC
 XX
 XX 11-APR-2001 (first entry)
 DT
 XX
 XX Corynebacterium glutamicum MCT protein SEQ ID NO:46.
 DE
 XX Corynebacterium glutamicum; brevbacterium lactofermentum; MCT;
 KW membrane construction and membrane transport protein; petroleum spill;
 KW hydrocarbon degradation; gram positive aerobic bacterium; marker;
 KW identification; microorganism; fine chemical production; transformation;
 KW genome mapping; genetic engineering.
 KW
 XX
 XX Corynebacterium glutamicum.
 OS
 XX

PN WO200100805-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000WO-IB00926.
 XX
 XX 25-JUN-1999; 99US-0141031.
 PR 08-JUL-1999; 99DE-1031454.
 PR 08-JUL-1999; 99DE-1031478.
 PR 08-JUL-1999; 99DE-1031563.
 PR 09-JUL-1999; 99DE-1032122.
 PR 09-JUL-1999; 99DE-1032124.
 PR 09-JUL-1999; 99DE-1032125.
 PR 09-JUL-1999; 99DE-1032128.
 PR 09-JUL-1999; 99DE-1032180.
 PR 09-JUL-1999; 99DE-1032182.
 PR 09-JUL-1999; 99DE-1032190.
 PR 09-JUL-1999; 99DE-1032209.
 PR 09-JUL-1999; 99DE-1032212.
 PR 09-JUL-1999; 99DE-1032227.
 PR 09-JUL-1999; 99DE-1032228.
 PR 09-JUL-1999; 99DE-1032229.
 PR 09-JUL-1999; 99DE-1032230.
 PR 14-JUL-1999; 99DE-1032927.
 PR 14-JUL-1999; 99DE-1033005.
 PR 14-JUL-1999; 99DE-1033006.
 PR 27-AUG-1999; 99DE-1040764.
 PR 27-AUG-1999; 99DE-1040765.
 PR 27-AUG-1999; 99DE-1040766.
 PR 27-AUG-1999; 99DE-1040830.
 PR 27-AUG-1999; 99DE-1040831.
 PR 27-AUG-1999; 99DE-1040832.
 PR 27-AUG-1999; 99DE-1040833.
 PR 31-AUG-1999; 99DE-1041378.
 PR 31-AUG-1999; 99DE-1041379.
 PR 31-AUG-1999; 99DE-1041395.
 PR 03-SEP-1999; 99DE-1042077.
 PR 03-SEP-1999; 99DE-1042078.
 PR 03-SEP-1999; 99DE-1042079.
 PR 03-SEP-1999; 99DE-1042088.
 XX
 PA (BAD) BASF AG.
 PI Pompejus M, Kroeger B, Schroeder H, Zeider O, Habernauer G;
 XX
 DR WPI; 2001-071486/08.
 DR N-PSDB; AAF67765.
 XX
 PT Corynebacterium glutamicum nucleic acids encoding membrane construction
 PT and membrane transport proteins or their portions, useful for typing or
 PT identifying C. glutamicum or related bacteria, and as markers for
 PT transformation -
 XX
 PS Claim 20; Page 202-207; 1119pp; English.
 CC AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
 CC construction and membrane transport (MCT) proteins given in AAB76510 to
 CC AAB76847. The MCT nucleic acids and proteins are useful in the
 CC identification of microorganisms which can be used to produce fine
 CC chemicals, for modulating fine chemical production in C. glutamicum or
 CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or
 CC identification of C. glutamicum or related bacteria, as reference points
 CC for mapping C. glutamicum genome, and as markers for transformation.
 CC AAF68082 and AAF68083 represent sequencing primers which are used in an
 CC example from the present invention.
 XX
 SQ Sequence 1874 AA;
 Query Match 4.1%; Score 9; DB 22; Length 1874;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 168 PAAEAPVAA 176
 |||||
 Db 568 PAAEAPVAA 576
 RESULT 12
 AAG90680
 ID AAG90680 standard; Protein; 2969 AA.
 XX
 AC AAG90680;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE C glutamicum protein fragment SEQ ID NO: 4434.
 XX
 KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KM organic acid synthesis.
 OS
 XX Corynebacterium glutamicum.
 PN EPI108790-A2.
 XX
 PD 20-JUN-2001.
 PF 18-DEC-2000; 2000EP-0127688.
 PR 16-DEC-1999; 99JP-0377484.
 PR 07-APR-2000; 2000JP-0159162.
 PR 03-AUG-2000; 2000JP-0280988.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 DR WPI; 2001-376931/40.
 DR N-PSDB; AAF65899.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT -
 XX
 PS Claim 17; SEQ ID NO: 4434; 246bp + Sequence Listing; English.
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 XX
 SQ Sequence 2969 AA;
 Query Match 4.1%; Score 9; DB 22; Length 2969;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 168 PAAEAPVAA 176
 |||||
 Db 1663 PAAEAPVAA 1671
 RESULT 13
 ABP96450
 ID ABP96450 standard; peptide; 9 AA.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139434.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150864.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

```
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

```
Query Match          3.7%; Score 8; DB 21; Length 65;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      130 AAEPA AAA 137
          |||||
Db       19 AAEPA AAA 26
```

Search completed: January 15, 2004, 15:28:31
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:23:11 ; Search time 18 Seconds

(without alignments)
572.158 Million cell updates/sec

Title: US-09-923-236-2

Perfect score: 219
Sequence: 1 MKLLIMACTIVCAVPAKRRF.....TAKPAAPHPSPSLEQANQ 219

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	9	4.1	322	1	RLA0 LEICH
2	9	4.1	323	1	RLA0 LEICH
3	9	4.1	581	1	HEM1 STRO
4	9	4.1	864	1	NIA VOLCA
5	8	3.7	107	1	RLA1 LEIPE
6	8	3.7	110	1	RLA3 SCHPO
7	8	3.7	128	1	RL7 CHLPN
8	8	3.7	353	1	CUP5 GALME
9	8	3.7	396	1	P53 MESAU
10	8	3.7	574	1	ATP2 CHLRE
11	8	3.7	575	1	MSPA TREMA
12	8	3.7	576	1	ECR HELVI
13	8	3.7	1477	1	KELC DROME
14	8	3.7	2470	1	NTC2 MOUSE
15	7	3.2	82	1	ANPA_PSEAM
16	7	3.2	101	1	RLI2 METTL
17	7	3.2	101	1	RLI2 METTL
18	7	3.2	102	1	RLI2 METUA
19	7	3.2	105	1	RLA2 DICDI
20	7	3.2	107	1	RLA1 CHLRE
21	7	3.2	109	1	RLA1 TRYCR
22	7	3.2	110	1	RLA2 CRYST
23	7	3.2	111	1	RA2A MAIZE
24	7	3.2	111	1	RLI2 AERPE
25	7	3.2	111	1	RLA1 CAEBL
26	7	3.2	111	1	RLA2 ASPFU
27	7	3.2	112	1	RLA1 DROME
28	7	3.2	112	1	RLI2 HALNI
29	7	3.2	114	1	RLA1 RAT
30	7	3.2	114	1	RLA2 EIMTE
31	7	3.2	120	1	CUI9 ARADI
32	7	3.2	125	1	PSAE SPIOL
33	7	3.2	129	1	DYLA_CHLRE

ALIGNMENTS

34	7	3.2	136	1	RL28 HUMAN	P46779 homo sapien
35	7	3.2	136	1	RL28 MOUSE	P41105 mus musculu
36	7	3.2	136	1	SR14_HUMAN	P37108 homo sapien
37	7	3.2	136	1	BCCP_PSEAE	P37799 pseudomonas
38	7	3.2	163	1	ATPD_CAEEL	Q09544 caenorhabd
39	7	3.2	168	1	ATPD_HUMAN	P30049 homo sapien
40	7	3.2	199	1	CYCY_RHOC	Q05389 rhodobacter
41	7	3.2	204	1	RS3A_METUA	O800F2 methanosarc
42	7	3.2	208	1	RS6 MYCGE	P47336 mycoplasma
43	7	3.2	215	1	EF1D_WHEAT	P29546 triticum ae
44	7	3.2	215	1	RS6 MYCPN	P75543 mycoplasma
45	7	3.2	220	1	COAT_CMVSI	Q00467 cymbidium m

RESULT 1

ID	RLA0 LEICH	STANDARD;	PRT;	322 AA.
AC	P39096;			
DT	01-FEB-1995 (Rel. 31, Created)			
DR	01-FEB-1995 (Rel. 31, Last sequence update)			
DE	01-OCT-1996 (Rel. 34, Last annotation update)			
GN	60S acidic ribosomal protein P0.			
OS	LCPO.			
OC	Leishmania chagasi.			
OX	Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.			
RN	NB1_TaxID=44271;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ISOLATE MHOM/BR/82/BA-2;			
RX	MEDLINE=9422252; PubMed=7513304;			
RA	Skelly Y.A.W., Benson D.R., Elwasli M., Badaro R., Burns J.M. Jr.,			
RA	Reed S.G.;			
RT	"Antigens shared by Leishmania species and Trypanosoma cruzi:			
RL	Immunol. 62:1643-1651 (1994)."			
CC	-1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT			
CC	OF E.COLI PROTEIN L10.			
CC	-1- SUBUNIT: P0 FORMS A PENTAMERIC COMPLEX BY INTERACTION WITH			
CC	DIMERS OF P1 AND P2.			
CC	-1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; L29300; AAA29263.1; .			
DR	InterPro; IPR001813; 60S_ribosomal.			
DR	InterPro; IPR001790; Ribosomal_L10.			
DR	Pfam; PF00428; 60S_ribosomal_1.			
DR	Pfam; PF00466; Ribosomal_L10; 1.			
KW	Ribosomal protein; Phosphorylation.			
SQ	SEQUENCE 322 AA; 34594 MW; 2389F125356D26D2 CRC64;			
Query Match	4.1%;	Score 9;	DB 1;	Length 322;
Best Local Similarity	100.0%;	Pred. No. 0.66;		
Matches	9;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;
OY	130 AAEPAAAP 138			
Db	289 AAEPAAAP 297			
RESULT 2				
ID	RLA0 LEICH	STANDARD;	PRT;	323 AA.
AC	P39097;			

```

DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 60S acidic ribosomal protein P0.
GN LIPO-A AND LIPO-B.
OS Leishmania infantum.
OC Eukaryota; Eukaryozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OK NCBI_Taxid=5671;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEM 75;
RA MEDLINE=94088674; PubMed=8264730;
RA Soto M., Requena J.M., Alonso C.;
RT "Isolation, characterization and analysis of the expression of the
RT Leishmania ribosomal P0 protein genes";
RL Mol. Biochem. Parasitol. 61:265-274(1993).
CC -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT
CC OF E.COLI PROTEIN L10.
CC -1- SUBUNIT: P0 FORMS A PENTAMERIC COMPLEX BY INTERACTION WITH
CC DIMERS OF P1 AND P2.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X72714; CAA51264.1; -
DR EMBL; X72714; CAA51263.1; -
DR InterPro; IPR001813; 60S_ribosomal.
DR InterPro; IPR001790; Ribosomal_L10.
DR Pfam; PF00428; 60S_ribosomal; 1.
DR Pfam; PF00466; Ribosomal_L10; 1.
KW Ribosomal protein; Phosphorylation.
SQ SEQUENCE 323 AA; 34771 MW; 14967BD7A439D69E CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 AAEPAAP 138
DB 290 AAEPAAP 298

RESULT 3
HEM1_STRCO STANDARD; PRT; 581 AA.
ID HEM1_STRCO
AC Q9WX15;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.1-) (GluTR).
GN HEMA OR SCO319 OR SCE68.17C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycetales; Streptomycetaceae; Streptomycetes.
OK NCBI_Taxid=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Laikre L., Murphy L., Oliver K., O'Neill S.,
RA Rabinovitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wierozorek A., Woodward J., Barrell B.G., Parkhill J.,

```

```

RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)";
RL Nature 417:141-147(2002).
CC -1- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + tRNA(Glu).
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC -1- SIMILARITY: belongs to the glutamyl-tRNA reductase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL939116; CAB45353.1; -
DR PIR; T36267; T36267.
DR HAMAP; MF_00087; atypical; 1.
DR InterPro; IPR000343; GluTR.
DR InterPro; IPR000594; ThLP_domain.
DR Pfam; PF00745; GluTR_dimer; 1.
DR Pfam; PF05201; GluTR_N; 1.
DR Pfam; PF05200; GluTR_NAD_bind; 1.
DR TIGRPFAM; TIGR01035; hemA; 1.
DR PROSITE; PS00747; GluTR; 1.
KW Porphyrin biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT DOMAIN 292 416
FT ACT SITE 50 50 NUCLEOPHILE (BY SIMILARITY).
FT ACT SITE 99 99 BASE (BY SIMILARITY).
SQ SEQUENCE 581 AA; 60562 MW; D4E256B105AFA037 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 VAAEPAA 162
DB 36 VAAEPAA 44

RESULT 4
NIA_VOLCA STANDARD; PRT; 864 AA.
ID NIA_VOLCA
AC P36841;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nitrate reductase [NADH] (EC 1.7.1.1) (NR).
GN NIT7A.
OS Volvox carterii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Volvocaceae; Volvox.
OK NCBI_Taxid=3067;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F. Nagatensis / HK10;
RA MEDLINE=93013022; PubMed=1398126;
RA Gruber H., Goetlink S.D., Kirk D.L., Schmitt R.;
RT "The nitrate reductase-encoding gene of Volvox carterii: map location,
RT sequence and induction kinetics.";
RL gene 120:75-83(1992).
CC -1- FUNCTION: Nitrate reductase is a key enzyme involved in the first
CC step of nitrate assimilation in plants, fungi and bacteria.
CC -1- CATALYTIC ACTIVITY: Nitrite + NAD(+) + H(2)O = nitrate + NADH.
CC -1- COFACTOR: EACH SUBUNIT OF THE ENZYME CONTAINS 1 EQUIVALENT OF FAD,
CC HEME IRON, AND MOLYBDENUM-PTERIN AS PROSTHETIC GROUPS. THE HEME
CC GROUP IS CALLED CYTOCHROME B-557.
CC -1- SUBUNIT: Homodimer (by similarity).
CC -1- INDUCTION: By nitrate.
CC -1- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
CC N-TERMINAL DOMAIN.

```

```

CC -1- SIMILARITY: Contains 1 cytochrome b5 heme-binding domain.
CC -1- SIMILARITY: TO PAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
CC C-TERMINAL DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: X64136; CAA45497.1; -.
CC PIR: J01432; J01422.
CC HSSP: P04166; 1B5M.
CC InterPro: IPR001199; Cyt_B5.
CC InterPro: IPR001834; Cyt_B5_reductase.
CC InterPro: IPR000572; Euk_Mb_oxred.
CC InterPro: IPR001709; FPM_cyt_redctse.
CC InterPro: IPR005066; Mo-co_dimer.
CC InterPro: IPR001433; Oxred_PAD/NAD(P).
CC Pfam: PF00970; FAD_binding_6; 1.
CC Pfam: PF00173; heme_1; 1.
CC Pfam: PF03404; Mo-co_dimer; 1.
CC Pfam: PF00175; NAD_binding_1; 1.
CC Pfam: PF00174; oxidored_molzb; 1.
CC PRINTS: PR00406; CYTB5RDTASE.
CC PRINTS: PR00363; CYTOCHROME_B5.
CC PRINTS: PR00407; EUMOPTERIN.
CC PRINTS: PR00371; EPNCR.
CC ProDom: PD00612; Cyt_B5; 1.
CC PROSITE: PS00191; CYTOCHROME_B5_1; 1.
CC PROSITE: PS0255; CYTOCHROME_B5_2; 1.
CC PROSITE: PS00559; MOLYBDOPTERIN_EUK; 1.
CC Oxidoreductase; Flavoprotein; FAD; NAD; Heme; Molybdenum;
CC Nitate assimilation.
CC METAL: 139 139 MOLYBDENUM-PTERIN (POTENTIAL).
CC METAL: 193 193 MOLYBDENUM-PTERIN (POTENTIAL).
CC DISULFID: 376 376 INTERCHALIN (POTENTIAL).
CC METAL: 532 532 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC METAL: 555 555 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC SEQUENCE 864 AA; 96402 MW; 499529652CDD1C7 CRC64;

Query Match 4.1%; Score 9; DB 1; Length 864;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 120 AAAAPAPP 128
Db 478 AAAAPAPP 486

RESULT 5
RLAI LEIPE STANDARD; PRT; 107 AA.
AC 04613;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S acidic ribosomal protein p1.
OS Leishmania peruviana.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5681;
RN NCB1_TaxID=5681;
RP SEQUENCE FROM N.A.
RC STRAIN=MHOM/PE/84/LC26;
RL de los Santos M., Carrillo C., Panebra A., Montoya Y.;
FU FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
CC PROTEIN SYNTHESIS.
CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
CC SUBUNIT.
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.

```

```

CC -1- SIMILARITY: Contains 1 cytochrome b5 heme-binding domain.
CC -1- SIMILARITY: TO PAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
CC C-TERMINAL DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AF045249; AAC02701.1; -.
CC InterPro: IPR001813; 60S_ribosomal.
CC InterPro: IPR001859; Ribosomal_P2.
CC Pfam: PF00428; 60S_ribosomal; 1.
CC PRINTS: PR00456; RIBOSOMALP2.
CC Ribosomal protein.
CC SEQUENCE 107 AA; 10896 MW; 9117DEC5E490A071 CRC64;

Query Match 3.7%; Score 8; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 119 SAAAPAA 126
Db 71 SAAAPAA 78

RESULT 6
RLA3 SCHPO STANDARD; PRT; 110 AA.
ID RLA3 SCHPO
AC P17477;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 60S acidic ribosomal protein p1-alpha 3 (A3).
GN RPA3 OR SPBC389.13C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN NCB1_TaxID=4896;
RP SEQUENCE FROM N.A.
RX MEDLINE=90220620; PubMed=2325655;
RA Beltrame M., Bianchi M.E.;
RT "A gene family for acidic ribosomal proteins in Schizosaccharomycetes
RT pombe: two essential and two nonessential genes.";
RL Mol. Cell. Biol. 10:2341-2348(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hitalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsle K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voiclaert G., Aert R., Robben J., Grymptre B.,
RA Welljens I., Vanstelele E., Rieger W., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer B., Moesli D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Bento J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

```

RA Cernuti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Usero D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
 CC PROTEIN SYNTHESIS.
 CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
 CC SUBUNIT.
 CC -1- MISCELLANEOUS: YEASTS CONTAIN 4 INDIVIDUAL SMALL RIBOSOMAL A
 CC PROTEINS (RPA) WHICH CAN BE CLASSIFIED INTO TWO COUPLES OF SIMILAR
 CC BUT NOT IDENTICAL SEQUENCES. EACH COUPLE IS DISTINCTLY RELATED TO
 CC ONE OF THE TWO A PROTEINS PRESENT IN MULTICELLULAR ORGANISMS.
 CC -1- MISCELLANEOUS: RPA3 AND RPA4 ARE ESSENTIAL FOR CELL SURVIVAL,
 CC WHEREAS RPA1 AND RPA2 ARE NOT.
 CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M3139; AAA35336.1; -;
 CC EMBL; AL023070; CA117793.1; -;
 CC PIR; C34715; R6BYP3. -;
 CC GeneDB; SP03B9.13c; -;
 CC DR InterPro; IPR001813; 608_ribosomal.1.
 CC Pfam; PF00428; 608_ribosomal.1.
 CC Ribosomal protein; Phosphorylation; Multigene family.
 CC SO SEQUENCE 110 AA; 11171 MW; 173AAC9779F5A91 CRC64;
 QY 119 SAAAPAA 126
 DB 64 SAAAPAA 71
 RT RESULT 7
 RT RT7_CHLNP STANDARD; PRT; 128 AA.
 ID RT7_CHLNP
 AC Q929A1; Q9JQ70;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 50S ribosomal protein L7/L12.
 GN RPL7 OR RLT7 OR CPN0080 OR CP0695.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiales; Chlamydiales; Chlamydiales.
 NCBI_Taxid=83558;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CWL029;
 RC MEDLINE=9920606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Greenwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RC MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,
 RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
 RA Liner K., Weisman J., Knout H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mofn and Chlamydia
 RT pneumoniae AR39.";

RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=0138;
 RC MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hatori M., Kohara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- FUNCTION: SEEMS TO BE THE BINDING SITE FOR SEVERAL OF THE FACTORS
 CC INVOLVED IN PROTEIN SYNTHESIS AND APPEARS TO BE ESSENTIAL FOR
 CC ACCURATE TRANSLATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE001593; AAD18233.1; -;
 CC EMBL; AE002228; AAF38503.1; -;
 CC EMBL; AP002545; BAA98290.1; -;
 CC PIR; C72122; C72122.
 CC PIR; H86500; H86500.
 CC DR HSPC; P02392; ICTF.
 CC DR PHCI-2DPAGE; Q929A1; -;
 CC TIGR; CP0695; -;
 CC HAMAP; MF 00368; -;
 CC InterPro; IPR000206; Ribosomal_L12.
 CC Pfam; PF00542; Ribosomal_L12; 1.
 CC DR Prodom; PD001326; Ribosomal_L12; 1.
 CC TIGRPFAM; TIGR00855; L12; 1.
 CC Ribosomal protein; Complete proteome.
 CC INIT MET 0 BY SIMILARITY.
 CC SO SEQUENCE 128 AA; 13461 MW; 4E2F17A5B057CC CRC64;
 QY 171 EAPVAAAP 178
 DB 49 EAPVAAAP 56
 RT RESULT 8
 RT RT5_GALME STANDARD; PRT; 353 AA.
 ID RT5_GALME
 AC Q24998;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Pupal cuticle protein PCP52 precursor (GMPCC52).
 GN PCP52.
 OS Galleria mellonella (Wax moth).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Pyralidae; Galleriinae; Galleria.
 NCBI_Taxid=7137;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=epidermis;
 RC MEDLINE=95291282; PubMed=7773255;
 RA Kolberg U., Obermaier B., Hirsch H., Kelber G., Wolbert P.;
 RT "Expression cloning and characterization of a pupal cuticle protein
 RT cDNA of Galleria mellonella L.";
 RL Insect Biochem. Mol. Biol. 25:355-363(1995).
 CC -1- FUNCTION: COMPONENT OF THE CUTICLE OF THE PUPA OF GALLERIA
 CC MELLONELLA.

```

CC -1- DEVELOPMENTAL STAGE: MAXIMAL EXPRESSION DURING THE FIRST DAY AFTER
CC PUPAL ECDYSIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X77514; CAA54650.1; -
DR PIR; S41958; S41958.
KW Structural protein; Cuticle; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 353 POLY-ALA.
FT DOMAIN 235 243
SQ SEQUENCE 353 AA; 35818 MW; C3A35B04EC049172 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 6.1; Length 353;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 119 SAAAPAA 126
Db 236 SAAAPAA 243

RESULT 9
P53_MESAU STANDARD; PRT; 396 AA.
ID P53_MESAU P97276;
AC 000366;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cellular tumor antigen p53 (Tumor suppressor p53).
GN TP53.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OC NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Syrian; TISSUE=Kidney;
RA MEDLINE=92210007; PubMed=1555773;
RX Legros Y., McIntyre P., Soussi T.;
RT "The CDNA cloning and immunological characterization of hamster p53.";
RL Gene 112:247-250(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Hou E.W., Wiseman R.;
RL Submitted (Apr-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Acts as a tumor suppressor in many tumor types; induces
CC growth arrest or apoptosis depending on the physiological
CC circumstances and cell type. Involved in cell cycle regulation as
CC a trans-activator that acts to negatively regulate cell division
CC by controlling a set of genes required for this process. One of
CC the activated genes is an inhibitor of cyclin-dependent kinases.
CC Apoptosis induction seems to be mediated either by stimulation of
CC Bax and Fas antigen expression, or by repression of Bcl-2
CC expression.
CC -1- SUBUNIT: Binds DNA as a homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DISEASE: p53 is found in increased amounts in a wide variety
CC of transformed cells. p53 is frequently mutated or inactivated
CC in many types of cancer.
CC -1- SIMILARITY: BELONGS TO THE P53 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M75144; AAA37085.1; -
DR EMBL; U07182; AAB41344.1; -
DR PIR; JH0633; JH0633.
DR HSSP; P04637; ITUP.
DR InterPro; IPR002117; P53.
DR Pfam; PF00870; P53; 1.
DR PRINTS; PR00386; P53SUPPRESSR.
DR PRODOM; PD002681; P53; 1.
DR PROSITE; PS00348; P53; 1.
KW Anti-oncogene; DNA-binding; Transcription regulation; Activator;
KW Nuclear protein; Phosphorylation; Apoptosis.
FT DOMAIN 1 45 TRANSCRIPTION ACTIVATION (ACIDIC).
FT DNA_BIND 105 295 BY SIMILARITY.
FT DOMAIN 328 359 OLIGOMERIZATION.
FT DOMAIN 371 390 BASIC (REPRESSION OF DNA-BINDING).
FT DOMAIN 314 326 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT MOD_RES 15 15 PHOSPHORYLATION (BY PRK) (BY
FT MOD_RES 395 395 SIMILARITY).
FT MOD_RES 188 188 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 395 395 G -> S (IN REF. 2).
FT CONFLICT 188 188 G -> S (IN REF. 2).
SQ SEQUENCE 396 AA; 43631 MW; 906EF0256809BE3 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 6.7; Length 396;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAP 127
Db 66 AAAAPAP 73

RESULT 10
ATP2_CHLRE STANDARD; PRT; 574 AA.
ID ATP2_CHLRE
AC P38482;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase beta chain, mitochondrial precursor (EC 3.6.3.14).
GN ATP2.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OC NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=9235386; PubMed=1386535;
RX Franzen L.-G., Falk G.;
RT "Nucleotide sequence of cDNA clones encoding the beta subunit of
RT mitochondrial ATP synthase from the green alga Chlamydomonas
RT reinhardtii: the precursor protein encoded by the cDNA contains both
RT an N-terminal presequence and a C-terminal extension.";
RL Plant Mol. Biol. 19:771-780(1992).
RN [2]
RP STRUCTURE BY NMR OF 1-26.
RA MEDLINE=9632639; PubMed=8706917;
RA Lancelotti J.-M., Gans P., Bouchayer E., Bally I., Arlaud G.J.,
RA Jacquot J.-P.;
RT "NMR structures of a mitochondrial transit peptide from the green
RT alga Chlamydomonas reinhardtii.";
RL FEBS Lett. 391:203-208(1996).
CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC GRADIENT ACROSS THE MEMBRANE. THE BETA CHAIN IS THE CATALYTIC
CC SUBUNIT.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)

```

```

CC      HAS THREE MAIN SUBUNITS: A, B AND C.
CC      -1- SUBCELLULAR LOCATION: Mitochondrial.
CC      -1- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; X61624; CAA43808.1; -.
CC      PIR; S23530; S23530.
CC      HSSP; P00829; 1BMF.
CC      InterPro; IPR003593; AAA_ATPase.
CC      InterPro; IPR000793; ATPase_a/DC.
CC      InterPro; IPR000194; ATPase_a/bcentre.
CC      InterPro; IPR004100; ATPase_a/bN.
CC      InterPro; IPR005722; F1_ATPase_beta.
CC      Pfam; PF00006; ATP-synt_ab; 1.
CC      Pfam; PF00306; ATP-synt_ab_C; 1.
CC      Pfam; PF02874; ATP-synt_ab_N; 1.
CC      SMART; SM00382; AAA; 1.
CC      TIGRFA; TIGR01039; atpd; 1.
CC      PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
CC      ATP synthetase; CF(1); Hydrogen ion transport;
CC      Hydrolyase; ATP-binding; Mitochondrion; Transl. peptide.
CC      TRANSIT 1 26 MITOCHONDRION (POTENTIAL).
CC      CHAIN 27 574 ATP SYNTHASE BETA CHAIN.
CC      NP BIND 183 190 ATP (BY SIMILARITY).
CC      SO SEQUENCE 574 AA; 61821 MW; 22B3C6C6D18FBCE CRC64;

Query Match 3.7%; Score 8; DB 1; Length 574;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 AAEPAAAA 137
DB 26 AAEPAAAA 33

RESULT 11
MSPA_TREMA STANDARD; PRT; 575 AA.
ID MSPA_TREMA
AC Q92413;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Major outer membrane protein mspa precursor (Major sheath protein).
GN MSPA.
OS Treponema maltophilum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=51160;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-34.
RC STRAIN=ATCC 51939 / BR;
RA MEDLINE=99121045; PubMed=9922270;
RA Heuner K., Choi B.-K., Schade R., Moter A., Otto A., Goebel U.B.;
RT "Cloning and characterization of a gene (mspa) encoding the major
RT sheath protein of Treponema maltophilum ATCC 51939 (T).";
RL J. Bacteriol. 181:1025-1029 (1999).
RN [2]
RP SUBCELLULAR LOCATION.
RA MEDLINE=2123882; PubMed=1131313;
RA Heuner K., Melzer U., Choi B.-K., Goebel U.B.;
RT "Outer sheath associated proteins of the oral spirochete Treponema
RT maltophilum.";
RL FEWS Microbiol. Lett. 197:187-193 (2001).
CC -1- FUNCTION: Major component of the outer membrane sheath.
CC -1- SUBCELLULAR LOCATION: Outer membrane-associated.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; Y17800; CAA76862.1; -.
CC      KW Outer membrane; Signal.
CC      FT SIGNAL 1 19
CC      CHAIN 20 575 MAJOR OUTER MEMBRANE PROTEIN MSPA.
CC      FT CHAIN 20 575
CC      SO SEQUENCE 575 AA; 62250 MW; C666B188F5DE16A0 CRC64;

Query Match 3.7%; Score 8; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 AAEPAAEA 162
DB 19 AAEPAAEA 26

RESULT 12
ECR_HELVI STANDARD; PRT; 576 AA.
ID ECR_HELVI
AC O18473;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ecdysone receptor (Ecdytectoid receptor) (20-hydroxy-ecdysone
DE receptor) (20E receptor) (HVECR).
GN ECR OR NR1H1.
OS Heliothis virescens (Noctuid moth) (Owllet moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Heliothinae; Heliothis.
OX NCBI_TaxID=7102;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRC;
RA MEDLINE=99457718; PubMed=10528411;
RA Martinez A., Scanlon D., Gross B., Petara S.C., Palli S.R.,
RA Greenland A.J., Windass J., Pongs O., Broad P., Jepson I.;
RT "Transcriptional activation of the cloned Heliothis virescens
RT (Lepidoptera) ecdysone receptor (HVECR) by muisterona.",
RL Insect Biochem. Mol. Biol. 29:915-930 (1999).
CC -1- FUNCTION: RECEPTOR FOR ECDYSONE. BINDS TO ECDYSONE RESPONSE
CC ELEMENTS (ECRES) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR1
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC      EMBL; Y09009; CAA70212.1; -.
CC      HSSP; P20393; 1AGY.
CC      InterPro; IPR000536; Hormone_rec_1lg.
CC      InterPro; IPR001723; Steroid_receptor.
CC      InterPro; IPR001628; ZnF_Csteroi.
CC      Pfam; PF00104; hormone_rec; 1.
CC      Pfam; PF00105; zf-C4; 1.
CC      PRINTS; PR00398; STRDHOMONER.
CC      PRINTS; PR00047; STROIDRINGER.
CC      PRODom; PD000035; ZnF_Csteroi. 1.
CC      SMART; SM00430; HOLI; 1.
CC      SMART; SM00399; ZnF_C4; 1.
CC      PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.

```


KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 FT Zinc-finger.
 FT DOMAIN 1 162 MODULATING (POTENTIAL).
 FT DNA BIND 163 228 NUCLEAR RECEPTOR-TYPE.
 FT ZN FING 163 183 C4-TYPE.
 FT ZN FING 199 223 C4-TYPE.
 FT DOMAIN 326 545 HORMONE-BINDING (POTENTIAL).
 SQ SEQUENCE 576 AA: 64638 MW: D13EF787BF263A8 CRC64;
 Query Match 3.7%; Score 8; DB 1; Length 576;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 140 TATPVAAE 147
 Db 552 TATPVAAE 559
 RESULT 13
 KELC_DROME STANDARD; PRT; 1477 AA.
 AC 004652; Q04653; Q9VDA2;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ring canal kelch protein [contains: Kelch short protein].
 GN KEL OR CG7210.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neuroptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 CC NCBI_Taxid=7227;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RP MEDLINE=93201592; PubMed=8453663;
 RA Xue F., Cooley L.;
 RT "Kelch encodes a component of intercellular bridges in Drosophila egg
 RT chambers.";
 RL Cell 72:681-693 (1993).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RP MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Aamatiides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champs M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Daveport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman I.D., Hernandez J.R., Houck J.,
 RA Hoeltin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalalet M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laekko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paulev J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 RN [3]
 RN CHARACTERIZATION.
 RP TISSUE=Embryo;
 RC MEDLINE=97236487; PubMed=9118811;
 RX Robinson D.N., Cooley L.;
 RA "Examination of the function of two kelch proteins generated by stop
 RT codon suppression.";
 RT Development 124:1405-1417 (1997).
 CC -1- FUNCTION: COMPONENT OF RING CANALS THAT REGULATES THE FLOW OF
 CC CYTOPLASM BETWEEN CELLS. MAY BE INVOLVED IN THE REGULATION OF
 CC CYTOPLASM FLOW FROM NURSE CELLS TO THE OOCYTE DURING OOGENESIS.
 CC BINDS ACTIN.
 CC -1- SUBCELLULAR LOCATION: INNER SURFACE OF CYTOPLASMIC BRIDGES OR RING
 CC CANALS PRESENT IN EGG CHAMBERS. SUBCORTICALLY IN IMAGINAL DISK
 CC EPITHELIA.
 CC -1- TISSUE SPECIFICITY: BOTH PROTEINS ARE EXPRESSED IN OVARIES, MALE
 CC TESTIS, OVARECTOMIZED FEMALES, CUTICLE, SALIVARY GLAND AND
 CC IMAGINAL DISKS. KELCH ORF1 IS THE PREDOMINANT PROTEIN AND IS ALSO
 CC EXPRESSED IN FAT BODIES. ON ENTRY INTO METAMORPHOSIS LEVELS OF
 CC FULL LENGTH PROTEIN INCREASE IN TESTIS AND IMAGINAL DISKS.
 CC -1- DEVELOPMENTAL STAGE: LARVAE, PUPAE AND ADULTS.
 CC -1- SIMILARITY: Contains 1 BTB/POZ domain.
 CC -1- SIMILARITY: Contains 6 Kelch repeats.
 CC -1- CAUTION: REP 3 BELIEVES RESIDUE 690 IS A SELENOCYSTEINE.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 CC EMBL, L08483; AAAS3471.1; -;
 CC EMBL, L08483; AAAS3472.2; -;
 CC EMBL, AE003657; AAF53651.1; AUT_SEQ.
 CC HSSP; 005516; 1CS3.
 CC FLYBase; FBgn0001301; kel.
 CC GO; GO:0007292; P:ogenesis; IMP.
 CC GO; GO:0007301; P:ring canal formation; IDA.
 CC InterPro; IPR000210; BTB_POZ.
 CC InterPro; IPR006651; Kelch.
 CC InterPro; IPR006652; Kelch_rep.
 CC Pfam; PF00651; BTB; 1.
 CC Pfam; PF01344; Kelch; 6.
 CC PRINTS; PR00501; KELCHREPEAT.
 CC SMART; SM00225; BTB; 1.
 CC SMART; SM00612; Kelch; 6.
 CC PROSITE; PS50097; BTB; 1.
 KW Cytochrome; Actin-binding; Selenium; Selenocysteine; Kelch repeat;
 KW Repeat.
 KM CHAIN 1 1477
 FT CHAIN 1 689
 FT DOMAIN 157 223
 FT REPEAT 404 449
 FT REPEAT 450 496
 FT REPEAT 498 543
 FT REPEAT 545 592
 FT REPEAT 594 639
 FT REPEAT 641 687
 FT DOMAIN 18 28
 FT DOMAIN 29 87
 KELCH PROTEIN.
 KELCH SHORT PROTEIN.
 BTB.
 KELCH 1.
 KELCH 2.
 KELCH 3.
 KELCH 4.
 KELCH 5.
 KELCH 6.
 ASN-RICH.
 GLN-RICH.

FT DOMAIN 29 36 POLY-GLN.
FT DOMAIN 78 83 POLY-GLN.
FT SE-CYS 690 690 PROBABLE.
FT CONFLICT 493 493 V -> A (IN REF. 1).
FT CONFLICT 596 596 A -> R (IN REF. 1).
FT CONFLICT 824 824 P -> L (IN REF. 1).
FT CONFLICT 858 858 G -> D (IN REF. 1).
FT CONFLICT 1083 1083 A -> R (IN REF. 1).
FT CONFLICT 1086 1086 A -> G (IN REF. 1).
SQ SEQUENCE 1477 AA; 160086 MW; 4851EAE9D9DBA47 CRC64;

Query Match 3.7%; Score 8; DB 1; Length 1477;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAP 127
Db 740 AAAAPAP 747

RESULT 14
NTC2 MOUSE
ID NTC2 MOUSE STANDARD; RT; 2470 AA.
AC 03516; 006008; 060941;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurogenic locus notch homolog protein 2 precursor (Notch 2) (Notch
B).
GN NOTCH2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Thymus;
RA Hamada Y., Higuchi M., Tsujimoto Y.;
RT "Complete amino acid sequence and multiform transcripts encoded by a
RT single copy of mouse Notch2 gene."
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 316-1518 FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Embryo;
RX MEDLINE=93178363; PubMed=8440332;
RA Lardelli M., Lendahl U.;
RT "Notch A and Notch B-two mouse Notch homologues coexpressed in a
RT wide variety of tissues."
RL Exp. Cell Res. 204:364-372(1993).
[3]
RP SEQUENCE OF 1765-2153 FROM N.A.
RX MEDLINE=97075110; PubMed=8917536;
RA Miller L.A., Bigas A., Kopan R., Brashem-Stein C., Bernstein I.D.,
RA Martin D.I.;
RT "Inhibition of granulocytic differentiation by mNotch1."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13014-13019(1996).
[4]
RP FUNCTION.
RX MEDLINE=99396706; PubMed=10393320;
RA Hamada Y., Kadokawa Y., Okabe M., Ikawa M., Coleman J.R.,
RA Tsujimoto Y.;
RT "Mutation in ankyrin repeats of the mouse Notch2 gene induces early
RT embryonic lethality."
RL Development 126:3415-3424(1999).
[5]
RP DEVELOPMENTAL STAGE, AND ALTERNATIVE SPLICING.
RX MEDLINE=95533893; PubMed=7609614;
RA Higuchi M., Kiyama H., Hayakawa T., Hamada Y., Tsujimoto Y.;
RT "Differential expression of Notch1 and Notch2 in developing and adult
RT mouse brain."
RL Brain Res. Mol. Brain Res. 29:263-272(1995).
[6]
RP POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS OF MET-1699.

RX MEDLINE=21523956; PubMed=11518718;
RA Saxena M.T., Schroeder E.H., Mumm J.S., Kopan R.;
RT "Murtine notch homologs (N1-4) undergo presenilin-dependent
RT proteolysis."
RL J. Biol. Chem. 276:40268-40273(2001).
[7]
RP POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS OF MET-1699.
RX MEDLINE=21374376; PubMed=11459941;
RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.;
RT "Conservation of the biochemical mechanisms of signal transduction
RT among mammalian Notch family members."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031(2001).
CC -1- FUNCTION: Functions as a receptor for membrane-bound ligands
CC Jagged1, Jagged2 and Delta1 to regulate cell-fate determination.
CC upon ligand activation through the released notch intracellular
CC domain (NICD) it forms a transcriptional activator complex with
CC RBP-J kappa and activates genes of the enhancer of split locus.
CC Affects the implementation of differentiation, proliferation and
CC apoptotic programs (By similarity). May play an essential role in
CC postimplantation development, probably in some aspect of cell
CC specification and/or differentiation.
CC -1- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
CC terminal fragment N(EC) which are probably linked by disulfide
CC bonds.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Following
CC proteolytical processing NICD is translocated to the nucleus.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O35516-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O35516-2; Sequence=VSP_001405;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Expressed in the brain, liver, kidney,
CC neuroepithelia, somites, optic vesicles and branchial arches, but
CC not heart.
CC -1- DEVELOPMENTAL STAGE: Expressed in the embryonic ventricular zone,
CC the postnatal ependymal cells, and the choroid plexus throughout
CC embryonic and postnatal development.
CC -1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
CC which is proteolytically cleaved by a furin-like convertase in the
CC trans-Golgi network before it reaches the plasma membrane to yield
CC an active, ligand-accessible form. Cleavage results in a C-
CC terminal fragment N(TM) and a N-terminal fragment N(EC). Following
CC ligand binding, it is cleaved by TNF-alpha converting enzyme
CC (TACE) to yield a membrane-associated intermediate fragment called
CC notch extracellular truncation (NEXT). This fragment is then
CC cleaved by presenilin dependent gamma-secretase to release a
CC notch-derived peptide containing the intracellular domain (NICD)
CC from the membrane.
CC -1- PTM: Phosphorylated.
CC -1- SIMILARITY: BELONGS TO THE NOTCH FAMILY.
CC -1- SIMILARITY: Contains 35 EGF-like domains.
CC -1- SIMILARITY: Contains 2 Lin/Notch repeats.
CC -1- SIMILARITY: Contains 6 ANK repeats.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sb.ch/announce/>
CC or send an email to license@isb-sb.ch).
CC -----
CC EMBL; D32210; BAA22094.1; -
CC EMBL; X68279; CAA48340.1; -
CC EMBL; U31881; AAC52924.1; -
CC PIR; A49175; A49175.
CC HSSP; P16109; IFSB.
CC MGD; MGI:97364; Notch2.
CC GO; GO:0005887; C:integral to plasma membrane; IC.
CC GO; GO:0005515; P:protein binding activity; IPI.
CC GO; GO:0002011; P:morphogenesis of an epithelial sheet; IMP.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:26:21 ; Search time 35 Seconds

(without alignments) 1614.672 Million cell updates/sec

Title: US-09-923-236-2

Perfect score: 219
Sequence: 1 MKLLIMACTIVCAFPARKRRF.....TAKPAAPRHPSPSLQANQ 219

Scoring table: OLIGO

Gapop 60.0 , Gapept 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacterioplasmid:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	4.1	73	5	Q8MSH9
2	9	4.1	75	5	Q8VBL6
3	9	4.1	344	11	Q9DSJ3
4	9	4.1	784	16	Q8PH14
5	9	4.1	841	5	Q9NEA5
6	9	4.1	2969	16	Q8NS46
7	8	3.7	125	10	Q8H4F8
8	8	3.7	139	16	Q8Y2D2
9	8	3.7	141	16	Q8Y2R3
10	8	3.7	145	10	Q9PFC0
11	8	3.7	145	10	Q8E516
12	8	3.7	193	10	Q9C7F7
13	8	3.7	193	10	Q8LE10
14	8	3.7	209	15	Q9DQU8
15	8	3.7	228	16	Q9A4A2
16	8	3.7	231	2	Q93ND2

17	8	3.7	233	2	Q8GCE5	Q8gce5 streptomyces
18	8	3.7	242	4	Q8NHG8	Q8nhg8 homo sapien
19	8	3.7	259	2	Q91UD4	Q91ud4 rhizobium m
20	8	3.7	262	10	Q8W2Q2	Q8w2q2 oryza sativ
21	8	3.7	266	16	Q8P710	Q8p710 xanthomonas
22	8	3.7	268	16	Q95192	Q95192 mycobacteri
23	8	3.7	272	10	Q8S2S9	Q8s2s9 thellungie
24	8	3.7	275	16	Q92WH2	Q92wh2 rhizobium m
25	8	3.7	285	11	Q8BHC2	Q8bhc2 mus musculu
26	8	3.7	291	10	Q40478	Q40478 nicotiana c
27	8	3.7	291	10	Q8H382	Q8h382 oryza sativ
28	8	3.7	302	5	Q9N9A5	Q9n9a5 leishmania
29	8	3.7	325	2	Q929X8	Q929x8 frauteuria s
30	8	3.7	336	16	Q9AB24	Q9ab24 caulobacter
31	8	3.7	362	11	Q9CX00	Q9cx00 mus musculu
32	8	3.7	364	16	Q8P885	Q8p885 xanthomonas
33	8	3.7	380	11	Q8B2V6	Q8b2v6 mus musculu
34	8	3.7	381	2	Q8E996	Q8e996 acinetobact
35	8	3.7	384	16	Q9AD02	Q9ad02 streptomyce
36	8	3.7	401	8	Q8WEA7	Q8wea7 primula jes
37	8	3.7	428	10	Q941Q8	Q941q8 zea mays (m
38	8	3.7	434	10	Q8H3E5	Q8h3e5 oryza sativ
39	8	3.7	443	8	Q8WEA0	Q8wea0 primula pol
40	8	3.7	448	8	Q31130	Q31130 streptitza
41	8	3.7	468	8	Q9GE00	Q9ge00 primula cor
42	8	3.7	468	8	Q8ME93	Q8me93 primula bul
43	8	3.7	493	2	Q84311	Q84311 klebsiella
44	8	3.7	493	2	Q9W4P4	Q9w4p4 drosophila
45	8	3.7	493	5	Q8S2R1	Q8s2r1 drosophila

ALIGNMENTS

RESULT 1	Q8MSH9	PRELIMINARY:	PRT:	73 AA.
ID	Q8MSH9	01-OCT-2002 (TREMBLrel. 22, Created)		
AC	Q8MSH9	01-OCT-2002 (TREMBLrel. 22, Last sequence update)		
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	GHI9893P (Fragment).			
DN	BCDNA:GHI9893			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkeley;			
RA	Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,			
RA	Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Fries E.,			
RA	George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,			
RA	Miranda A., Mungall C.J., Nunoo J., Paclele J., Paragas V., Park S.,			
RA	Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,			
RA	Celniker S.;			
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AY118802; AAM5062.1; -			
DR	PIRBASE: P89063235; BCDNA:GHI9893.			
FT	NON_TER			
SQ	SEQUENCE	73 AA; 6847 MW; BC8B77C976694F2D CRC64;		
Qy	Query Match	4.1% Score 9; DB 5; Length 73;		
Db	Best Local Similarity	100.0%; Pred. No. 1.1;		
	Matches	9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
		120 AAAAPAP 128		
		43 AAAAPAP 51		
RESULT 2				

Q9VBL6 PRELIMINARY; PRT; 75 AA.
 AC Q9VBL6; Q24390;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Accessory gland-specific peptide 57da precursor (Male accessory gland secretory protein 57da).
 GN M5757DA OR CG9074.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RP STRAIN=BERKELEY;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazer V., Champs M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abiri J.F., Agayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhakov S., Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P., Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Dwyer A.D., Dew I., Dietz S.M., Dodson K.J., Evans J.E., Gabor C., Gabor C., Ferrera S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris J., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C., Jaiswal M., Kalish F., Karpen G.H., Ke Z., Kemnison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mallett B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stepieton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasmann D.A., Weinstein G.M., Weisenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: THIS PROTEIN IS TRANSFERRED FROM MALE TO FEMALE DURING MATING AND MAY AFFECT EGGLAYING AND BEHAVIOR AFTER MATING.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: LUMEN FLUID OF MALE ACCESSORY GLANDS, BECOMES SEMINAL FLUID.
 CC EMBL; Z33647; CAA83925.1; -;
 CC EMBL; A5003753; AAF56515.1; -;
 DR FlyBase; FBgn0011668; M5757da.
 KW Signal; Behavior.

FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 75 ACCESSORY GLAND-SPECIFIC PEPTIDE 57DA.
 FT DOMAIN 39 64 ALA/PRO-RICH DOMAIN.
 FT CONFLICT 39 46 MISSING (IN REF. 1).
 FT CONFLICT 64 75 MISSING (IN REF. 1).
 SQ SEQUENCE 75 AA; 7106 MW; BCD3D70817C98E14 CRC64;
 Query Match 4.1%; Score 9; DB 5; Length 75;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 120 AAAAPAPP 128
 DB 45 AAAAPAPP 53
 RESULT 3
 ID Q9D5J3 PRELIMINARY; PRT; 344 AA.
 AC Q9D5J3;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE 4930432K09Rik protein.
 GN 4930432K09Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Testis;
 RC MEDLINE=2108560; PubMed=11217851;
 RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa K., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S., Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K., Saito T., Okazaki Y., Gotohori T., Bono H., Katsukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T., Fleschmann W., Gaasterland T., Giesi C., King B., Koehne H., Kuenl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schirml L.M., Stabli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gostincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H., Lyons P., Marchionni L., Mashima U., Mazzarelli U., Komatsu N., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S., Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK015291; BAB29782.1; -;
 DR MGD; MGI:191029; 4930432K09Rik.
 DR InterPro; IPR002965; P rich exten.
 DR PRINTS; PR01217; PRICHTENSEN.
 SQ SEQUENCE 344 AA; 36294 MW; D6EAD71CE5802AC CRC64;
 Query Match 4.1%; Score 9; DB 11; Length 344;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGLLMACI 9
 DB 1 MGLLMACI 9
 RESULT 4
 ID Q8PH14 PRELIMINARY; PRT; 784 AA.
 AC Q8PH14;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)

```

DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Hypothetical protein XAC3446.
GN XAC3446.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cammaran F., Cardoso J., Chamberg F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorxy H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Tindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Sebuel J.C., Kitaajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463 (2002).
DR EMBL; AF011991; AAM38289.1; -
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF00486; trans_reg_C_1.
DR ProDom; PD000329; Trans_reg_C_1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 764 AA; 84093 MW; F41377005382EC8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 16; Length 784;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 APVAAEPAA 160
DB 160 APVAAEPAA 168

RESULT 5
Q9NEA5 PRELIMINARY; PRT; 841 AA.
ID Q9NEA5;
AC Q9NEA5;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Possible inner dynein arm.
GN L5515.03.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Masny D., Fumelle B., Goffeau A., Ivens A.C., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=9814635; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145 (1998).
DR EMBL; AL161414; CAB77677.1; -

```

```

DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 4.
DR SMART; SM00320; WD40; 4.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 841 AA; 93125 MW; BA1B72F1BE821226 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 5; Length 841;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 SAAPAPAP 127
DB 40 SAAPAPAP 48

RESULT 6
Q8NS46 PRELIMINARY; PRT; 2969 AA.
ID Q8NS46;
AC Q8NS46;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE 3-oxoacyl-(acyl-carrier-protein) synthase (EC 2.3.1.85).
GN CGJ0836.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005276; BAB98229.1; -
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR003965; Fatty_acid_synth.
DR InterPro; IPR000794; Ketoacyl-synth.
DR InterPro; IPR002539; Maoc_dehydratase.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF00698; Acyl_transf_1.
DR Pfam; PF00109; ketoacyl-synth_1.
DR Pfam; PF02801; ketoacyl-synth_C_1.
DR Pfam; PF01575; Maoc_dehydratase; 1.
DR PRINTS; PR01483; FASYNTHASE.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR PROSITE; PS00435; PEROXIDASE_1; 1.
KW Transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 2969 AA; 31512 MW; 908E454660D90815 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 16; Length 2969;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 PAAPAPVAA 176
DB 1663 PAAPAPVAA 1671

RESULT 7
Q8H4F8 PRELIMINARY; PRT; 125 AA.
ID Q8H4F8;
AC Q8H4F8;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE OJ1221 H04.16 protein.
GN OJ1221 H04.16.
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```

```

OC  Eubacteriaceae; Oryzae; Oryza.
OK  NCBI_TaxID=39947;
RN  (1)
RP  SEQUENCE FROM N.A.
RC  STRAIN=CV. Nipponbare;
RA  Sasaki T., Matsunoto T., Yamamoto K.;
RT  "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
    clone:O1121.H04."
RL  Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF004183; BAC21432.1; -.
SQ  SEQUENCE 125 AA; 13560 MW; 31F375D8398B8A52 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  121 AAAPAAP 128
    |||||
    11 AAAPAAP 18

RESULT 8
OBY2D2  PRELIMINARY; PRT; 139 AA.
ID  OBY2D2:
AC  OBY2D2:
DT  01-MAR-2002 (TrEMBLrel. 20, Created)
DT  01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT  01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE  Hypothetical signal peptide protein RSC0404.
GN  RSC0404 OR RSC03372.
OS  Ralstonia solanacearum (Pseudomonas solanacearum).
OC  Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
    Ralstoniaceae; Ralstonia.
OX  NCBI_TaxID=305;
RN  (1)
RP  SEQUENCE FROM N.A.
RC  STRAIN=CM11000;
RX  MEDLINE=21681879; PubMed=11823852;
RA  Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
    Arlat M., Billault A., Broctier P., Camus J.C., Catolico L.,
    Chandler M., Choienne N., Claudel-Renard C., Cunac S., Demange N.,
    Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
    Stigter P., Thebault P., Whalen M., Winkler P., Levy M.,
    Weisenbach J., Boucher C.A.;
RT  "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL  Nature 415:497-502 (2002).
DR  EMBL; AL646059; CAD13932.1; -.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 139 AA; 14153 MW; A31350B6FD629039 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  120 AAAPAAP 127
    |||||
    55 AAAPAAP 62

RESULT 9
OBY2R3  PRELIMINARY; PRT; 141 AA.
ID  OBY2R3:
AC  OBY2R3:
DT  01-MAR-2002 (TrEMBLrel. 20, Created)
DT  01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT  01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE  Hypothetical protein RSC0272.
GN  RSC0272 OR RSC0239.
OS  Ralstonia solanacearum (Pseudomonas solanacearum).
OC  Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
    Ralstoniaceae; Ralstonia.
OX  NCBI_TaxID=305;
RN  (1)

```

```

RP  SEQUENCE FROM N.A.
RC  STRAIN=CM11000;
RX  MEDLINE=21681879; PubMed=11823852;
RA  Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
    Arlat M., Billault A., Broctier P., Camus J.C., Catolico L.,
    Chandler M., Choienne N., Claudel-Renard C., Cunac S., Demange N.,
    Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
    Stigter P., Thebault P., Whalen M., Winkler P., Levy M.,
    Weisenbach J., Boucher C.A.;
RT  "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL  Nature 415:497-502 (2002).
DR  EMBL; AL646058; CAD13800.1; -.
DR  InterPro; IPR001763; Rhodanese-like.
DR  SMART; SM00450; RHOD; 1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 141 AA; 15454 MW; 878FE1840A8B8B4E CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  149 AAGAPVAA 156
    |||||
    126 AAGAPVAA 133

RESULT 10
O9PFC0  PRELIMINARY; PRT; 145 AA.
ID  O9PFC0:
AC  O9PFC0:
DT  01-MAR-2001 (TrEMBLrel. 16, Created)
DT  01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT  01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE  Histone H2B like protein.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
    OC  eustoid II; Brassicales; Brassicaceae; Arabidopsids.
OX  NCBI_TaxID=3702;
RN  (1)
RP  SEQUENCE FROM N.A.
RC  STRAIN=Columbia;
RX  MEDLINE=97471969; PubMed=9330910;
RA  Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
    Miyajima N., Tabata S.;
RT  "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
    features of the 1.6 Mb regions covered by twenty physically assigned
    P1 clones.";
RL  DNA Res. 4:215-230 (1997).
KW  -1- SUBUNIT: THE NUCLEOSOME IS AN OCTAMER CONTAINING TWO MOLECULES OF
    H2A, H2B, H3, AND H4; WHICH WRAP APPROXIMATELY 146 BP OF DNA (BY
    SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC  CC -1- SIMILARITY: BELONGS TO THE HISTONE H2B FAMILY.
DR  EMBL; AB005243; BAB10609.1; -.
DR  InterPro; IPR004822; Histone core.
DR  InterPro; IPR000558; Histone_H2B.
DR  Pfam; PF00125; histone_1.
DR  PRINTS; PR00621; HISTONEH2B.
DR  ProDom; PD000497; Histone_H2B; 1.
DR  SMART; SM00427; H2B; 1.
DR  PROSITE; PS00357; HISTONE_H2B; 1.
KW  Chromosomal protein; DNA-binding; Nuclear protein; Nucleosome core.
SQ  SEQUENCE 145 AA; 15732 MW; C68421B559D42B30 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  130 AAAPAAA 137
    |||||
    19 AAAPAAA 26

```



```

RESULT 11
ID 096516 PRELIMINARY; PRT; 145 AA.
AC 096516;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Histone H2B like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
RX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ecotype columbia;
RA Philids G.;
RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
CC -1- SUBUNIT: THE NUCLEOSOME IS AN OCTAMER CONTAINING TWO MOLECULES OF
CC H2A, H2B, H3, AND H4; WHICH WRAP APPROXIMATELY 146 BP OF DNA (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: Y07745; CAA69025.1; -
DR InterPro: IPR004822; Histone_core.
DR InterPro: IPR000558; Histone_H2B.
DR Pfam: PF00125; histone_1.
DR PRINTS: PR00621; HISTONEH2B.
DR ProDom: PD000497; HISTONE_H2B.
DR SMART: SM00427; H2B.1.
DR PROSITE: PS00357; HISTONE_H2B; 1.
DR Chromosomal protein; DNA-binding; Nuclear protein; Nucleosome core.
SQ SEQUENCE 145 AA; 15733 MW; 6C8EB8B18390FA686 CRC64;

```

```

Query Match 3.7%; Score 8; DB 10; Length 145;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 130 AAAPAAA 137
Db 19 AAAPAAA 26

```

```

RESULT 12
ID 09C7F7 PRELIMINARY; PRT; 193 AA.
AC 09C7F7;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Lipid transfer protein, putative.
GN F1K9.6 OR ATIG27950.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
RX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA MEDLINE=21016719; PubMed=1130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Alatafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Chauk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Feng B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Mafti R., Marziani A.,
RA Miltsecher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,

```

```

RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakono H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
RA Uterback T., Van Aken S., Vayenberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakono H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakono H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC069471; AAG51485.1; -
DR EMBL: AY092956; AAM12955.1; -
DR EMBL: AY128712; AAM91112.1; -
DR InterPro: IPR003612; AA1.
DR Pfam: PF00234; tryp_alpha_amy1.1.
SQ SEQUENCE 193 AA; 19759 MW; D54B38B12FEB610 CRC64;

```

```

Query Match 3.7%; Score 8; DB 10; Length 193;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 120 AAAPAAP 127
Db 20 AAAPAAP 27

```

```

RESULT 13
ID 08LE10 PRELIMINARY; PRT; 193 AA.
AC 08LE10;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Lipid transfer protein, putative.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
RX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzer S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation."
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana."
RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AY085407; AAM62634.1; -
DR InterPro: IPR003612; AA1.
DR Pfam: PF00234; tryp_alpha_amy1.1.
DR SMART: SM00499; AA1.1.
SQ SEQUENCE 193 AA; 19747 MW; C091074205D76610 CRC64;

```

Query Match 3.7%; Score 8; DB 10; Length 193;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAP 127
 |||||
 Db 20 AAAAPAP 27

RESULT 14

Q9DQ08 PRELIMINARY; PRT; 209 AA.

AC Q9DQ08: 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
 DE Nef protein (Negative factor) (27 kDa protein).
 GN NEF.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_Taxid=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=01u09;
 RX MEDLINE=21002575; PubMed=1118071;
 RA Geflin R., Wolf D., Muller R., Hill M.D., Steilweg E., Freitag M.,
 Sasse G., Scott G.B., Baur A.S.;
 RT "Functional and structural defects in HIV-1 nef genes derived from
 pediatric long-term survivors.";
 RL AIDS Res. Hum. Retroviruses 16:1855-1868(2000).
 CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
 ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
 SIMILARITY).
 CC EMBL; AF252901; AAG34573.1; -.
 DR HSSP; P03406; IEFN.
 DR InterPro; IPR001558; HIV_Nef.
 DR Pfam; PF00469; F-protein_1.
 DR ProDom; PD000031; HIV_Nef_1.
 KW AIDS; GTP-binding; Lipoprotein; Myristate.
 SQ SEQUENCE 209 AA; 23656 MW; 6B858BAD98C79519 CRC64;

Query Match 3.7%; Score 8; DB 15; Length 209;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 AEPAPAP 138
 |||||
 Db 23 AEPAPAP 30

RESULT 15

Q9A4A2 PRELIMINARY; PRT; 228 AA.

AC Q9A4A2: 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Cytochrome c, membrane-bound.
 GN CC2935.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_Taxid=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nieman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 Bisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 Kolonay J.F., Smit J., Craven M.B., Kouri H., Shetty J., Berry K.,

RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE005958; AAK24897.1; -.
 DR HSSP; P81459; 1154.
 DR TIGR; CC2935; -.
 DR InterPro; IPR000104; Antifreeze 1.
 DR InterPro; IPR000345; Cyt heme_bind.
 DR InterPro; IPR003088; Cyt_Ct.
 DR InterPro; IPR002327; Cyt_CTAB.
 DR Pfam; PF00034; Cytochrome_c; 1.
 DR PRINTS; PR00308; ANTIFREEZE1.
 DR PRINTS; PR00604; CYTCHROME_CTAB.
 DR ProDom; PD000375; Cyt_CTAB; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; 1.
 KW Complete proteome.
 SQ SEQUENCE 228 AA; 23023 MW; 5AE05C7DD5D05F82 CRC64;

Query Match 3.7%; Score 8; DB 16; Length 228;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAP 127
 |||||
 Db 219 AAAAPAP 226

Search completed: January 15, 2004, 15:29:49
 Job time : 35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:26:56 ; Search time 21 Seconds
(without alignments)
1002.902 Million cell updates/sec

Title: US-09-923-236-2

Perfect score: 219
Sequence: 1 MKLLIMACTVCVAFARKRF.....TAKPAAPRHPSPSLQANQ 219

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : PIR 76:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	4.1	581	2 T36267	probable glutamyl-nitrate reductase
2	9	4.1	864	1 JCI422	60S acidic ribosomal L7/L12 ribosomal protein
3	8	3.7	110	1 R6BYP3	probable lipid transfer protein
4	8	3.7	129	2 H86500	cytochrome c, membrane
5	8	3.7	129	2 C72122	probable ABC transporter
6	8	3.7	193	2 H86404	probable ABC transporter
7	8	3.7	228	2 E87612	cytochrome c, membrane
8	8	3.7	268	2 G70645	hypothetical protein
9	8	3.7	276	2 G95887	hypothetical protein
10	8	3.7	291	2 T02434	DNA binding protein
11	8	3.7	325	2 T46873	electron transfer
12	8	3.7	336	2 A87300	TPR domain protein
13	8	3.7	353	2 S41958	pupal cuticle protein
14	8	3.7	381	2 T46827	phenoxylbenzoate diester
15	8	3.7	396	1 JH0633	cellular tumor antigen
16	8	3.7	495	2 T52066	probable farnesyl transferase
17	8	3.7	506	2 A40679	transcription enhancer
18	8	3.7	523	2 B40679	transcription enhancer
19	8	3.7	550	2 T75557	hypothetical protein
20	8	3.7	574	2 S23530	H+-transporting ATPase
21	8	3.7	611	2 C96030	probable cell division
22	8	3.7	846	2 S52418	GRP-binding protein
23	8	3.7	1476	2 A45773	kelch protein, long
24	8	3.7	4307	2 T20721	hypothetical protein
25	7	3.2	38	2 S23173	phosphorylase kinase
26	7	3.2	78	2 E83446	hypothetical protein
27	7	3.2	82	1 FDFLAW	antifreeze protein
28	7	3.2	82	2 S02326	antifreeze protein
29	7	3.2	82	2 A05161	antifreeze protein

30	7	3.2	82	2 I51125	antifreeze protein
31	7	3.2	100	2 F95852	hypothetical protein
32	7	3.2	101	2 H69091	ribosomal protein
33	7	3.2	102	2 D64363	ribosomal protein
34	7	3.2	106	1 R6DOP2	acidic ribosomal protein
35	7	3.2	107	1 R6KMLC	acidic ribosomal protein
36	7	3.2	109	1 R6UTP1	acidic ribosomal protein
37	7	3.2	110	2 T37490	ribosomal protein
38	7	3.2	111	2 E72524	probable ribosomal protein
39	7	3.2	111	2 B95857	hypothetical protein
40	7	3.2	112	1 R5FFZE	acidic ribosomal protein
41	7	3.2	112	2 S54179	acidic ribosomal protein
42	7	3.2	112	2 E86141	protein T25K16.9
43	7	3.2	114	1 R5RT12	acidic ribosomal protein
44	7	3.2	114	1 R5HS2H	ribosomal protein
45	7	3.2	114	2 F84266	50S ribosomal protein

ALIGNMENTS

RESULT 1
T36267
Probable glutamyl-tRNA reductase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36267
R:Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, June 1999
A:Reference number: 221576
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-581 <MDR>
A:Cross-references: EMBL:AL079345; PIDN:CAB4533.1; GSPDB:GN00070; SCOPDB:SCE68.17C
A:Experimental source: strain AJ(2)
A:Genetics:
A:Gene: SCOPDB:SCE68.17C

Query Match 4.1%; Score 9; DB 2; Length 581;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 VAAEPAA 162
DB 36 VAAEPAA 44

RESULT 2
JCI422
Nitrate reductase (NADH) (EC 1.7.1.1) - Volvox carterii
C:Species: Volvox carterii
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: JCI422; S22192
R:Gruber, H.; Goetinck, S.D.; Kirk, D.L.; Schmitt, R.
Gene 120, 75-83, 1992
A:Title: The nitrate reductase-encoding gene of Volvox carterii: Map location, sequence
A:Reference number: JCI422; MUID:93013022; PMID:1398126
A:Accession: JCI422
A:Molecule type: DNA
A:Residues: 1-864 <GR2>
A:Cross-references: EMBL:X64136; NID:g21993; PIDN:CAA45497.1; PID:g21994
A:Note: Submitted to the EMBL Data Library, January 1992
C:Genetics:
A:Gene: nita
A:Map position: linkage group IX
A:Insertions: 183/3; 234/2; 295/1; 339/1; 372/1; 425/2; 521/3; 593/3; 677/3; 797/2
C:Function: catalyzes the reduction of nitrate to nitrite
A:Description: catalyzes the reduction of nitrate to nitrite
C:Superfamily: nitrate reductase (NADH); cytochrome b5 core homology; cytochrome-b5 reductase
C:Keywords: dimer; electron transfer; PAD; flavoprotein; heme; iron; metalloprotein; mol
F:36-425/Domain: molybdopterin-binding domain homology <PCO>
F:497-571/Domain: cytochrome b5 core homology <CB5>

F:613-864/Domain: cytochrome-b5 reductase homology <CBR>
 F:139/Binding site: molybdopterin (Cys) (covalent) #status predicted
 F:532,555/Binding site: heme iron (His) (axial ligands) #status predicted

Query Match 4.1%; Score 9; DB 1; Length 864;
 Best Local Similarity 100.0%; Pred. No. 2.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAAP 128
 DB 478 AAAAPAAP 486

RESULT 3

60s acidic ribosomal protein p1-alpha - fission yeast (Schizosaccharomyces pombe)
 N:Alternate names: acidic ribosomal protein 3; ribosomal protein YL2e11
 C:Species: Schizosaccharomyces pombe
 C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 10-Dec-1999
 C:Accession: C34715; T40352
 R:Beltrame, M.; Bianchi, M.E.
 M:1. Cell. Biol. 10, 2341-2348, 1990
 A:Title: A gene family for acidic ribosomal proteins in Schizosaccharomyces pombe: two
 A:Reference number: A34715; MUID:90220620; PMID:2325655
 A:Accession: C34715

A:Molecule type: DNA
 A:Residues: 1-110 <BEL>
 A:Cross-references: EMBL:A33339; NID:g173467; PTDN:AAA5336.1; PID:g173468
 R:Wood, V.; Rajandream, M.A.; Barrett, B.G.; Skelton, J.; Churcher, C.M.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z21922

A:Accession: T40352
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-110 <WOO>
 A:Cross-references: EMBL:AL022070; PTDN:CAA17793.1; GSPDB:GN00067; SPDB:SPBC3B9.13c
 A:Experimental source: strain 972h-; cosmid c3B9
 C:Genetics:

A:Gene: rps3
 A:Map position: 2
 A:introns: 23/3; 103/3
 C:Superfamily: rat acidic ribosomal protein P1
 C:Keywords: phosphoprotein; protein biosynthesis; ribosome

Query Match 3.7%; Score 8; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 SAAAPAA 126
 DB 64 SAAAPAA 71

RESULT 4

H86500
 L7/L12 ribosomal protein [imported] - Chlamydia pneumoniae (strain J138)
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: H86500
 R:Shitai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349; PMID:10871362
 A:Accession: H86500

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-129 <STO>
 A:Cross-references: GB:BA000008; NID:g8978453; PTDN:BA98290.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: r17
 C:Superfamily: Escherichia coli ribosomal protein L12

Query Match 3.7%; Score 8; DB 2; Length 129;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 EAPVAAEP 178
 DB 50 EAPVAAEP 57

RESULT 5

C72122
 ribosomal protein L7/L12 CP0695 [imported] - Chlamydia pneumoniae (strains CWL029 an
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
 C:Accession: C72122; G81548
 R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
 Nature Genet. 21, 385-389, 1999
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:99206606; PMID:10192388
 A:Accession: C72122

A:Molecule type: DNA
 A:Residues: 1-129 <ARN>
 A:Cross-references: GB:AE001593; GB:AE001363; NID:g4376334; PTDN:AAD18233.1; PID:g437633
 A:Experimental source: strain CWL029
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of Chlamydia trachomatis Morn and Chlamydia pneumoniae AR39.
 A:Reference number: A81500; MUID:20150255; PMID:10684935
 A:Accession: G81548

A:Molecule type: DNA
 A:Residues: 1-129 <REA>
 A:Cross-references: GB:AE002228; GB:AE002161; NID:g7189606; PTDN:AAF8503.1; PID:g718960
 A:Experimental source: strain AR39, HL cells
 C:Genetics:
 A:Gene: r17; CP0695
 C:Superfamily: Escherichia coli ribosomal protein L12

Query Match 3.7%; Score 8; DB 2; Length 129;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 EAPVAAEP 178
 DB 50 EAPVAAEP 57

RESULT 6

H86404
 probable lipid transfer protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: H86404
 R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.;
 Jensen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shum, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yi, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: H86404

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-193 <STO>
 A:Cross-references: GB:AE005172; NID:g11024871; PTDN:AAG26955.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 3.7%; Score 8; DB 2; Length 193;

Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAP 127
Db 20 AAAAPAP 27

RESULT 7

cytochrome c, membrane-bound [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 18-Nov-2002

C:Accession: E87612

R:Finan, M.C.; Feldlym, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Land, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwim, M.L.; Haft, D.H.; Kolon

n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:1125647

A:Accession: E87612

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-228 <STO>

A:Cross-references: GB:AB005673; NID:g13424561; PIDN:AAK24897.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC2935

C:Superfamily: membrane-bound cytochrome c/cytochrome c homology

C:Keywords: chromoprotein; heme; iron; metalloprotein

F:81.84/Binding site: heme (Cys) (covalent) #status predicted

F:85/Binding site: heme iron (His) (axial ligand) #status predicted

F:150/Binding site: heme iron (Met) (axial ligand) #status predicted

Query Match 3.7%; Score 8; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 AAAAPAP 127
Db 219 AAAAPAP 226

RESULT 8

hypothetical protein RV3134c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: G70645

R:Coile, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: G70645

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-268 <COL>

A:Cross-references: GB:Z83867; GB:AL123456; NID:g3261695; PIDN:CAB06280.1; PID:e230951;

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV3134c

Query Match 3.7%; Score 8; DB 2; Length 268;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 PSPAPAT 200
Db 138 PSPAPAT 145

RESULT 9

G95887

probable ABC transporter permease protein Smb20381 (imported) - Sinorhizobium meliloti

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 01-Mar-2002

C:Accession: G95887

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: G95887

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-276 <KRP>

A:Cross-references: GB:AL591985; PIDN:CAC48767.1; PID:g15140240; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid pSymB

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barilay-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: Smb20381

A:Genome: plasmid

C:Superfamily: spermidine/putrescine transport system permease protein potI

Query Match 3.7%; Score 8; DB 2; Length 276;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 AAEPAAA 137
Db 268 AAEPAAA 275

RESULT 10

T02434

DNA binding protein EREBP-4 - common tobacco

C:Species: Nicotiana tabacum (common tobacco)

C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 31-Mar-2001

C:Accession: T02434

R:Ohme-Takagi, M.; Shinshi, H.

Plant Cell 7, 173-182, 1995

A:Title: Ethylene-inducible DNA binding proteins that interact with an ethylene response

A:Reference number: Z14671; MUID:95276459; PMID:7756828

A:Accession: T02434

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-291 <OHM>

A:Cross-references: EMBL:D38125; NID:G790361; PIDN:BA07323.1; PID:g1208497

A:Experimental source: strain BY4; tissue-type leaf

C:Superfamily: Arabidopsis thaliana hypothetical protein T19P19.170

Query Match 3.7%; Score 8; DB 2; Length 291;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 133 PAAAPLT 140
Db 247 PAAAPLT 254

RESULT 11

T48873

electron transfer protein [imported] - Frateuria sp. (strain ANA-18)

C:Species: Frateuria sp.

A:Variety: strain ANA-18

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jun-2000

C:Accession: T48873

R.Murakami, S.; Takashima, A.; Takemoto, J.; Takenaka, S.; Shinke, R.; Aoki, K.

Gene 226, 189-198, 1999

A>Title: Cloning and sequence analysis of two catechol-degrading gene clusters from the

A/Reference number: Z24833

A/Accession: T46873

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-325 <MUR>

A/Cross-references: EMBL:AB009373; PIDN:BAW75213.1

A/Experimental source: strain ANA-18

C/Superfamily: phthalate dioxygenase reductase; cytochrome-b5 reductase homology; ferred

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 325;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 AAEPAAA 137
Db 223 AAEPAAA 230

RESULT 12

TPR domain protein [imported] - Caulobacter crescentus

C/Species: Caulobacter crescentus

C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C/Accession: A87300

R/Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A>Title: Complete Genome Sequence of Caulobacter crescentus.

A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: A87300

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-336 <STO>

A/Cross-references: GB:AE005673; NID:gl3421571; PIDN:AAK22397.1; GSPDB:GN00148

C/Genetics:

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 336;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAPAP 127
Db 112 AAAPAP 119

RESULT 13

S41958

pupal cuticle protein precursor - greater wax moth

C/Species: Galleria mellonella (greater wax moth)

C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 19-May-2000

C/Accession: S41958

R/Kollberg, U.; Obermeier, B.; Hirsch, H.; Kelber, G.; Wolbert, P.

submitted to the EMBL Data Library, February 1994

A/Description: Expression cloning and characterization of a pupal cuticle protein cDNA G

A/Reference number: S41958

A/Accession: S41958

A/Molecule type: mRNA

A/Residues: 1-353 <KOL>

A/Cross-references: EMBL:X77514; NID:g453403; PIDN:CAA54650.1; PID:g453404

F/1-15/Domain: signal sequence #status predicted <SIG>

F/16-353/Product: pupal cuticle protein #status predicted <MAT>

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 353;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 SAAAPAA 126
|||||||

Db 236 SAAAPAA 243

RESULT 14

T46827

phenoxylbenzoate dioxygenase [imported] - Acinetobacter lwoffi

C/Species: Acinetobacter lwoffi

C/Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 02-Sep-2000

C/Accession: T46827

R/Kim, S.I.; Leem, S.H.; Choi, J.S.; Chung, Y.H.; Kim, S.; Park, Y.M.; Ha, K.S.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z24100

A/Accession: T46827

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-381 <KIM>

A/Cross-references: EMBL:U77659; PIDN:AAC31770.1

A/Experimental source: strain K24

C/Superfamily: phthalate dioxygenase reductase; cytochrome-b5 reductase homology; ferred

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 381;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 AAEPAAA 137
Db 279 AAEPAAA 286

RESULT 15

JH0633

cellular tumor antigen p53 - golden hamster

N/Alternate names: tumor-suppressor protein p53

C/Species: Mesocricetus auratus (golden hamster)

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C/Accession: JH0633

R/Legros, Y.; McIntyre, P.; Soussi, T.

Gene 112, 247-250, 1992

A>Title: The cDNA cloning and immunological characterization of hamster p53.

A/Reference number: JH0633; MUID:9221007; PMID:1555773

A/Accession: JH0633

A/Molecule type: mRNA

A/Residues: 1-396 <LEG>

A/Cross-references: GB:M75144; NID:G191414; PIDN:AA37085.1; PID:G191415

A/Experimental source: kidney, strain MPI

C/Genetics:

A/Supersfamily: cellular tumor antigen p53

C/Keywords: apoptosis; cell division control; DNA binding; homotetramer; nucleus; phosp

F/179,182,241,245/Binding site: zinc (Cys, His, Cys, Cys) #status predicted

F/395/Binding site: phosphoryl-RNA (Ser) (covalent) #status predicted

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 396;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAPAP 127
Db 66 AAAPAP 73

Search completed: January 15, 2004, 15:30:23
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:07:29 ; Search time 41 Seconds

(without alignments)
847.832 Million cell updates/sec

Title: US-09-923-236-2.

Perfect score: 1178

Sequence: 1 MKLLLMACIVCAFAFKRRF.....TAKPAPRPHSPSPLEQANQ 219

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

A.GeneSeq.19Jun03:*

- 1: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
- 2: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
- 3: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
- 4: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
- 5: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
- 6: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
- 7: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
- 8: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
- 9: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
- 10: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
- 11: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
- 12: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
- 13: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
- 14: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
- 15: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
- 16: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
- 17: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
- 18: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
- 19: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
- 20: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
- 21: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
- 22: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
- 23: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*
- 24: /SIDS1/gcsgdata/geneSeq/geneSeq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1178	100.0	219	22	AA660109
2	1178	100.0	219	23	AA627862
3	1178	100.0	219	23	ABG31608
4	1178	100.0	219	23	AAU74536
5	1178	100.0	219	24	ABO08515
6	1172	99.5	219	20	AAW30653
7	1113.5	94.5	221	20	AAV19472
8	183.5	15.6	325	22	ABG21919
9	180	15.3	207	14	AA40797

10	179	15.2	207	17	AAW06915
11	179	15.2	207	20	AAV23850
12	179	15.2	207	20	AAV23305
13	175.5	14.9	718	22	AAW42089
14	175.5	14.9	903	23	ABP41277
15	175.5	14.9	903	23	AAW40303
16	174	14.8	566	22	ABW61040
17	173	14.7	538	22	AAW82806
18	173	14.7	1427	23	ABP63024
19	171.5	14.6	693	23	ABP65929
20	170.5	14.5	149	22	AAU0451
21	170	14.4	267	19	AAW81726
22	170	14.4	267	19	AAW63559
23	170	14.4	267	20	AAV39156
24	170	14.4	267	20	AAV39013
25	170	14.4	428	23	AAW50963
26	170	14.4	428	24	ABP70532
27	170	14.4	842	22	ABW66631
28	170	14.4	864	22	ABW71319
29	170	14.4	1028	24	ABU70689
30	170	14.4	3151	24	ABR39571
31	170	14.4	3173	23	ABW90743
32	170	14.4	3173	24	ABW54450
33	170	14.4	3176	24	ABR39570
34	169.5	14.4	783	19	AAW37151
35	169.5	14.4	787	19	AAW37152
36	169.5	14.4	802	22	AAU09139
37	169.5	14.4	802	22	ABW63276
38	169	14.3	299	22	ABW82807
39	168.5	14.3	550	22	ABW66693
40	168.5	14.3	706	22	ABW66693
41	165.5	14.0	763	18	AAW31852
42	164	13.9	572	17	AAW31855
43	163.5	13.8	214	18	AAW65913
44	162.5	13.8	471	22	ABW61785
45	162	13.8	1064	22	ABW66662

ALIGNMENTS

RESULT 1	AA660109	standard; Protein; 219 AA.
AC	AA660109;	
XX		
DT	28-MAR-2001 (first entry)	
XX		
DE	Human transport protein TPPT-29.	
XX		
KW	Human; transport protein; TPPT; transport disorder; metabolic disorder;	
KW	neurological disorder; cardiovascular disorder; reproductive disorder;	
KW	immune disorder; cancer.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200078953-A2.	
XX		
PD	28-DEC-2000.	
XX		
PF	16-JUN-2000; 2000MO-US16668.	
XX		
PR	17-JUN-1999; 99US-0139923.	
PR	10-AUG-1999; 99US-0148177.	
PR	18-AUG-1999; 99US-0149357.	
PR	28-OCT-1999; 99US-0162287.	
XX		
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
PI	Lai P, Yang J, Yue H, Hallman JL, Tang YT, Bandman O, Burford N;	
PI	Baughn MR, Aizawa Y, Lu DM, Au-Young J, Patterson C;	

T. cruzi Tcd anti g
Tcd protein sequen
Tcd antigen of Trp
Human polypeptide
Human ovarian anti
Human polypeptide
Drosophila melanog
Human low density
Human polypeptide
Human polypeptide
Protein encoded by
M. tuberculosis Im
Mycobacterium tube
M. tuberculosis an
M. tuberculosis re
Maize methyl Cpg b
Histone deacetylase
Drosophila melanog
Drosophila melanog
Human adipocyte Se
Human alpha3 colla
Human Tumour Endoc
Human tumour endoc
Human alpha3 colla
Mouse neural Mena+
Mouse neural Mena+
Mammalian enabled
Drosophila melanog
Rabbit low density
Drosophila melanog
Mycobacterium tube
Mycobacterium tube
Cotton fiber-speci
Drosophila melanog
Drosophila melanog

DR WPI: 2001-041424/05.
 DR N-PSDB; AAF27729.
 XX Isolated polypeptide with a human transport protein sequence is useful
 PT for the diagnosis, prevention and treatment of disorders associated
 PT with the immune, reproductive and cardiovascular systems -
 XX
 PS Claim 2; Page 130; 165pp; English.
 XX
 CC The present invention provides the protein and coding sequences for 43
 CC novel human transport proteins (designated rPTs). These can be used in
 CC the diagnosis and treatment of transport, metabolic, neurological,
 CC reproductive, cardiovascular and immune disorders, and cell proliferative
 CC disorders such as cancer.
 XX
 SQ Sequence 219 AA;
 Query Match 100.0%; Score 1178; DB 22; Length 219;
 Best Local Similarity 100.0%; Pred. No. 2.5e-85;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKLLIACIVCAFAAKRRPPFAGEDNDGHLHPSLNIPIGIRNLPPPLYRPNVTVP 60
 DB 1 MKLLIACIVCAFAAKRRPPFAGEDNDGHLHPSLNIPIGIRNLPPPLYRPNVTVP 60
 QY 61 SYPGNTYTDTG.PSYWMIITSGFPYVHIRGPIATOLNVPPLPRGFPVPSRFSFA 120
 DB 61 SYPGNTYTDTG.PSYWMIITSGFPYVHIRGPIATOLNVPPLPRGFPVPSRFSFA 120
 QY 121 AAAPAPPIAAEPAAAPLTATPVAAPAAAGAVAAEPAAEAAPVGAEPAAEAAPVAAEPAA 180
 DB 121 AAAPAPPIAAEPAAAPLTATPVAAPAAAGAVAAEPAAEAAPVGAEPAAEAAPVAAEPAA 180
 QY 181 EAPVGEPAAEPSPAEPATAKPADEPHPSLEQANQ 219
 DB 181 EAPVGEPAAEPSPAEPATAKPADEPHPSLEQANQ 219
 RESULT 2
 AAE27862
 ID .AAE27862 standard; Protein; 219 AA.
 AC AAE27862;
 DT 27-DEC-2002 (first entry)
 DE Human zsig63 protein.
 DE Human zsig63 protein.
 KW Human; secreted salivary protein; zsig63 protein; host defense protein;
 KW immune modulating factor; antipathogenic; cell-cell signalling molecule;
 KW growth factor; cytokine; growth factor hormone activity; dental carrier;
 KW infection; tooth decay; periodontal disease; gastrointestinal disease;
 KW thrush; urinary tract infection; vaginal infection; diabetes; obesity;
 KW anti-inflammatory; chronic tissue damage; lung dysfunction; restenosis;
 KW gene therapy; salivary gland dysfunction; prostate gland dysfunction;
 KW forensic DNA profiling; chondrosarcoma; atherosclerosis; chromosome 4.
 XX
 OS Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT 1..15
 FT Peptide /label= Signal_peptide
 FT Protein 16..219 /note= "Mature human zsig63 protein"
 FT Region 14..21
 FT /note= "Hydrophilic region"
 FT Domain 16..37
 FT /note= "Domain 1"
 FT Region 17..33
 FT /note= "Antigenic epitope"
 FT Region 24..30
 FT /note= "Hydrophilic region"
 FT Domain 38..126

FT /note= "Domain 2"
 FT 66..73
 FT /note= "Antigenic epitope"
 FT Region 103..108
 FT /note= "Antigenic epitope"
 FT Region 124..133
 FT /note= "Repeat 1"
 FT 127..219
 FT /note= "Domain 3"
 FT Region 134..138
 FT /note= "Repeat 2"
 FT Region 139..143
 FT /note= "Repeat 3"
 FT Region 144..148
 FT /note= "Repeat 4"
 FT 149..153
 FT /note= "Repeat 5"
 FT Region 154..158
 FT /note= "Repeat 6"
 FT Region 159..163
 FT /note= "Repeat 7"
 FT 164..168
 FT /note= "Repeat 8"
 FT 169..173
 FT /note= "Repeat 9"
 FT Region 174..178
 FT /note= "Repeat 10"
 FT Region 179..183
 FT /note= "Repeat 11"
 FT 184..188
 FT /note= "Repeat 12"
 FT 187..192
 FT /note= "Hydrophilic region"
 FT Region 189..193
 FT /note= "Repeat 13"
 FT 190..197
 FT /note= "Antigenic epitope"
 FT Region 194..198
 FT /note= "Repeat 14"
 FT 199..203
 FT /note= "Repeat 15"
 FT Region 202..215
 FT /note= "Antigenic epitope"
 FT Region 204..208
 FT /note= "Repeat 16"
 FT
 XX US2002090677-A1.
 XX
 PD 11-JUL-2002.
 XX
 PF 03-AUG-2001; 2001US-0923236.
 XX
 XX 17-MAR-1999; 99US-124820P.
 PR 17-MAR-2000; 2000US-0527345.
 XX
 PA (ADLER/) ADLER D A.
 PA (SHEP/) SHEPPARD P O.
 XX
 PI Adler DA, Sheppard PO;
 XX
 DR WPI: 2002-642378/69.
 DR N-PSDB; AAD45050.
 XX
 PT Novel secreted salivary polypeptide, zsig63, useful as antimicrobial
 PT agent for treating microbial infection, dental carrier, periodontal
 PT disease, thrush gastrointestinal disease, and for aiding digestion -
 XX
 PS Claim 10; Page 28; 33pp; English.
 XX
 CC The invention relates to human secreted salivary polypeptide designated
 CC as zsig63 and nucleic acid molecules encoding such polypeptides. zsig63
 CC can be used in detecting agonists and antagonists of its activity, and
 CC is also useful as a host defense polypeptide, immune modulating factor,

CC antipathogenic polypeptide, cell-cell signalling molecule, growth factor,
 CC cytokine, or as secreted extracellular matrix associated proteins with
 CC growth factor hormone activity. It is useful for treating conditions
 CC associated with pathological microbes, including bacterial, fungal and
 CC viral infections, for treating dental caries (tooth decay), periodontal
 CC disease, thrush and gastrointestinal disease, for treating urinary tract
 CC infection, vaginal infection and for preventing infection in skin and
 CC other epithelial wounds, zsig63 is useful for establishing normal
 CC microflora and protect against pathogenic colonisation and invasion, for
 CC treating chronic tissue damage e.g. damage in extremities associated
 CC with diabetes and useful as anti-inflammatory agents. It is useful as a
 CC marker of lung dysfunction, salivary gland dysfunction, or dysfunction of
 CC prostate gland. It is also therapeutically useful for aiding digestion.
 CC Polynucleotides of the invention are used in gene therapy for increasing
 CC or inhibiting zsig63 activity, for detecting abnormalities on human
 CC chromosome 4 associated with disease or other human traits and as
 CC diagnostics in forensic DNA profiling. Sequences of the invention are
 CC useful for stimulating proliferation or differentiation of cardiac
 CC myocytes, for promoting proliferation or differentiation of adipocytes and for
 CC inhibiting chondrosarcomas, atherosclerosis, restenosis and obesity.
 CC The present sequence is human zsig63 protein. zsig63 gene is located
 CC on chromosome 4.

SQ Sequence 219 AA;

Query Match 100.0%; Score 1178; DB 23; Length 219;
 Best Local Similarity 100.0%; Pred. No. 2.5e-85;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLLIMACIVCAVFAFKRRFPFIGEDDNDGHLPLPSILNIPYGINLPPPLYRPPVNTVP 60
 DB 1 MKLLIMACIVCAVFAFKRRFPFIGEDDNDGHLPLPSILNIPYGINLPPPLYRPPVNTVP 60
 QY 61 SYPGNTYTDGTPSYPMWITSGPFVYHIRGFPPLATQINVPPLPRGPFVPPSRFESA 120
 DB 61 SYPGNTYTDGTPSYPMWITSGPFVYHIRGFPPLATQINVPPLPRGPFVPPSRFESA 120
 QY 121 AAAPAPAPAPAPAPAPATATPVAAEPAAAGVAAEPAAEPVGAEPAAEPVAAEPAA 180
 DB 121 AAAPAPAPAPAPAPAPATATPVAAEPAAAGVAAEPAAEPVGAEPAAEPVAAEPAA 180
 QY 181 EAPVGEPAAPAEPSPAEPATAKPAAPDEPHSPSLQANQ 219
 DB 181 EAPVGEPAAPAEPSPAEPATAKPAAPDEPHSPSLQANQ 219

RESULT 3

ABG31608 ID ABG31608 standard; Protein; 219 AA.

AC ABG31608;

DT 15-NOV-2002 (first entry)

DE Human secreted salivary protein zsig63.

KM Human, secreted salivary protein; zsig63; immunogen; zsig63-cytokine;
 KM antibody-cytokine; in vivo killing; pathological microbe; bacteria;
 KM fungal; viral; infection; salivary gland; anti-microbial; dental caries;
 KM tooth decay; periodontal disease; thrush; gastrointestinal disease;
 KM urinary tract infection; vaginal infection; skin infection; microflora;
 KM epithelial wound; pathogenic colonisation; invasion; pro-inflammatory;
 KM chronic tissue damage; vascular system; diabetes; anti-inflammatory;
 KM incompetent immune system; AIDS; acquired immunodeficiency syndrome;
 KM chemotherapy; radiation treatment; lung infection; cystic fibrosis;
 KM digestion; chromosome 4.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Region 1..15 /note= "Fusion protein peptide, specifically claimed
 FT in claim 18"

FT Region 1..219 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 9..204 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 14..19 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 16..21 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 16..37 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 16..219 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 17..33 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 24..29 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 24..33 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 25..30 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 38..126 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 66..73 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 103..108 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 109..215 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 124..133 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 127..219 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 134..138 /label= Repeat_2
 FT 139..143 /label= Repeat_3
 FT 144..148 /label= Repeat_4
 FT 149..153 /label= Repeat_5
 FT 154..158 /label= Repeat_6
 FT 159..163 /label= Repeat_7
 FT 164..168 /label= Repeat_8
 FT 169..173 /label= Repeat_9
 FT 174..178 /label= Repeat_10
 FT 179..183 /label= Repeat_11
 FT 184..188 /label= Repeat_12
 FT 187..192 /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT 189..193

FT FT /label= Repeat_13
 FT Region 190..197
 FT /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 194..198
 FT /label= Repeat_14
 FT Region 199..203
 FT /label= Repeat_15
 FT /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 202..215
 FT /label= Repeat_16
 FT /note= "Antigenic peptide, specifically claimed in
 FT claim 15"
 FT Region 204..208
 FT /label= Repeat_16
 FT US2002081701-A1.
 PN 27-JUN-2002.
 PD 03-AUG-2001; 2001US-0922480.
 XX 17-MAR-1999; 99US-124820P.
 XX 17-MAR-2000; 2000US-0527345.
 PR (ADLER/) ADLER D A.
 PA (SHEP/) SHEPPARD P O.
 XX
 XX Adler DA, Sheppard PO;
 PI WPI; 2002-635468/68.
 XX N-PSDB; ABS52633, ABS52634.
 DR
 DR Novel secreted salivary protein, zsig63 and polynucleotide encoding it
 PT useful for treating microbial infections, inflammatory conditions,
 PT dental caries and lung infections associated with cystic fibrosis
 XX
 PS Claim 10; Page 28; 33pp; English.
 XX
 CC The present invention relates to a new secreted salivary protein, zsig63.
 CC The invention is useful for detecting, in a test sample, the presence of
 CC an antagonist or agonist of zsig63 protein activity. The invention is
 CC also useful as an immunogen for producing an antibody to zsig63
 CC polypeptide. zsig63-cytokine fusion proteins or antibody-cytokine fusion
 CC protein are useful for enhancing in vivo killing of target tissues.
 CC Pharmaceutical composition comprising purified zsig63 polypeptide are
 CC useful in the treatment of conditions associated with pathological
 CC microbes, including bacterial, fungal and viral infections. High
 CC expression of zsig63 in salivary gland suggests that anti-microbial
 CC polypeptides are useful for treatment of dental caries (tooth decay),
 CC periodontal disease, thrush and gastrointestinal disease. Other
 CC applications can be used in urinary tract infections, vaginal infections,
 CC prevention of infection in skin and other epithelial wounds. The
 CC polypeptides can be used to establish normal microflora and protect
 CC against pathogenic colonisation and invasion. The invention is useful
 CC when pro-inflammatory activity is desired. Applications for
 CC such pro-inflammatory activity include the treatment of chronic tissue
 CC damage, particularly in areas having a limited or damaged vascular system
 CC e.g., damage in extremities associated with diabetes. Antagonists to
 CC zsig63 polypeptides may be useful as anti-inflammatory agents. The
 CC invention is useful for the treatment of patients having incompetent
 CC immune system, such as AIDS (acquired immunodeficiency syndrome) patients
 CC or individuals that have undergone chemotherapy, radiation treatment. The
 CC invention is also useful for the treatment of lung infections associated
 CC with cystic fibrosis and its agonists or antagonists are useful for
 CC aiding digestion. The present amino acid sequence represents the human
 CC secreted salivary protein zsig63 of the invention. This sequence is
 CC encoded by the human zsig63 gene located on chromosome 4.
 XX
 SQ Sequence 219 AA;
 Query Match 100.0%; Score 1178; DB 23; Length 219;
 Best Local Similarity 100.0%; Pred. No. 2.5e-85;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLLMACIVCAVAFARKRFPPIGEDNDGDGHPLPSINIPYGINRLPEPLYRRPVNTVP 60
 |||||
 Db 1 MGLLMACIVCAVAFARKRFPPIGEDNDGDGHPLPSINIPYGINRLPEPLYRRPVNTVP 60
 QY 61 SYRGNTYTDYGLPSYPMWILTSBPFPYVYHIRGFIPLATOLNVPPLPRGFPYPSRFFSA 120
 |||||
 Db 61 SYRGNTYTDYGLPSYPMWILTSBPFPYVYHIRGFIPLATOLNVPPLPRGFPYPSRFFSA 120
 QY 121 AAAPAAPPIAAEPAPAAAELTATPVAAEPAAAGAPVAAEPAAEPVGAEPAAEPVAAEPAA 180
 |||||
 Db 121 AAAPAAPPIAAEPAPAAAELTATPVAAEPAAAGAPVAAEPAAEPVGAEPAAEPVAAEPAA 180
 QY 181 EAPGVGEPAABEPPSPAPBPATKAPAPBPPSPSLQANQ 219
 |||||
 Db 181 EAPGVGEPAABEPPSPAPBPATKAPAPBPPSPSLQANQ 219
 RESULT 4
 AAU74536
 ID AAU74536 strand; Protein; 219 AA.
 XX
 XX AAU74536;
 AC
 XX
 XX 23-APR-2002 (first entry)
 DT
 XX
 XX Human zsig63 polypeptide.
 DE
 XX Human; zsig63; chromosome 4q12-q13; salivary protein; antimicrobial;
 XX microbial infection; tooth decay; periodontal disease; thrush; emphysema;
 XX gastrointestinal disease; urinary tract infection; vaginal infection;
 XX skin infection; epithelial wound; chronic tissue damage; cystic fibrosis;
 XX acquired immunodeficiency syndrome; AIDS; lung infection; sarcoidosis;
 XX chronic bronchitis; gene therapy; protein therapy.
 KW
 OS Homo sapiens.
 XX
 XX US6331413-B1.
 PN 18-DEC-2001.
 PD 17-MAR-2000; 2000US-0527345.
 XX 17-MAR-1999; 99US-124820P.
 XX PR
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 PI Adler DA, Sheppard PO;
 XX
 XX WPI; 2002-096707/13.
 DR N-PSDB; AAS20591.
 XX
 PT Polynucleotides encoding salivary proteins useful as anti-microbial
 PT agents -
 XX
 PS Claim 1; Column 49-52; 29pp; English.
 XX
 CC The invention relates to a polynucleotide derived from the 4q12-4q13
 CC region of human chromosome 4 and encoding a zsig63 polypeptide, a
 CC secreted salivary protein with anti-microbial activity. Due to their
 CC microbial activity, the sequences can be used in the study of microbial
 CC infections, e.g. for recombinant production of anti-microbial proteins.
 CC The sequences can be used in the treatment of tooth decay, periodontal
 CC disease, thrush, gastrointestinal disease, urinary tract infections,
 CC vaginal infections, skin infections, epithelial wounds, chronic tissue
 CC damage, acquired immunodeficiency syndrome (AIDS), cystic fibrosis, lung
 CC infections, sarcoidosis, emphysema and chronic bronchitis. This sequence
 CC represents human zsig63.
 XX
 SQ Sequence 219 AA;
 Query Match 100.0%; Score 1178; DB 23; Length 219;
 Best Local Similarity 100.0%; Pred. No. 2.5e-85;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MKLLIMACTIVCAAPARKRRPFPIGEDNDNDGHPHPSNITVYGRINLPPLIYRPNVTP	60	FT	Peptide	187..192	/label= Immunogenic_peptide	FT
Db	1	MKLLIMACTIVCAAPARKRRPFPIGEDNDNDGHPHPSNITVYGRINLPPLIYRPNVTP	60	FT	Peptide	187..192	/label= Immunogenic_peptide	FT
Qy	61	SYPGNTYTDGLPSYPMILITSGPFYVYHIGRFPPLATOLANVPPLPRGFPVPVPSRPFSA	120	FT	Peptide	24..33	/label= Immunogenic_peptide	FT
Db	61	SYPGNTYTDGLPSYPMILITSGPFYVYHIGRFPPLATOLANVPPLPRGFPVPVPSRPFSA	120	FT	Peptide	24..33	/label= Immunogenic_peptide	FT
Qy	121	AAAPAAPPIAAEPAAADLTATPVAAEPAAAGAPVAAEPAAAPVGAEPAAAPVAAEPAA	180	FT	Peptide	17..33	/label= Immunogenic_peptide	FT
Db	121	AAAPAAPPIAAEPAAADLTATPVAAEPAAAGAPVAAEPAAAPVGAEPAAAPVAAEPAA	180	FT	Peptide	17..33	/label= Immunogenic_peptide	FT
Qy	181	EAPVGEPAAPSPAPAPATAKPAAPBPSPSPLEQANO	219	FT	Peptide	103..108	/label= Immunogenic_peptide	FT
Db	181	EAPVGEPAAPSPAPAPATAKPAAPBPSPSPLEQANO	219	FT	Peptide	103..108	/label= Immunogenic_peptide	FT
RESULT 5	ID	ABU08515 standard; Protein; 219 AA.		FT	Peptide	190..197	/label= Immunogenic_peptide	FT
ABU08515				FT	Peptide	190..197	/label= Immunogenic_peptide	FT
AC	ABU08515;			FT	Peptide	190..197	/label= Immunogenic_peptide	FT
XX	28-MAY-2003 (first entry)			FT	Peptide	202..215	/label= Immunogenic_peptide	FT
XX				FT	Peptide	190..215	/label= Immunogenic_peptide	FT
XX	Human zsig63 protein.			FT	Peptide	124..133	/label= Immunogenic_peptide	FT
XX				FT	Peptide	134..138	/label= Immunogenic_peptide	FT
XX	Human; immunogen; zsig63; adhesin; salivary gland; dental carries; peridontal disease; thrush; gastrointestinal disease; epithelial wound; urinary tract infection; vaginal infection; skin infection; pro-inflammatory; chronic tissue damage; vascular system; diabetes; AIDS; lung infection; cystic fibrosis; lung dysfunction; digestive; salivary gland carcinoma; Pneumocystis carinii infection; emphysema; chronic bronchitis; prostate dysfunction; prostate adenocarcinoma; cell culture media; gene therapy; human chromosome 4q12-4q13; dentinogenesis imperfecta; dentin dysplasia type II.			FT	Peptide	149..153	/label= Immunogenic_peptide	FT
XX				FT	Peptide	154..158	/label= Immunogenic_peptide	FT
XX				FT	Peptide	159..163	/label= Immunogenic_peptide	FT
XX	Homo sapiens.			FT	Peptide	164..168	/label= Immunogenic_peptide	FT
XX				FT	Peptide	169..173	/label= Immunogenic_peptide	FT
XX	Key			FT	Peptide	174..178	/label= Immunogenic_peptide	FT
XX	Peptide			FT	Peptide	179..183	/label= Immunogenic_peptide	FT
XX				FT	Peptide	184..188	/label= Immunogenic_peptide	FT
XX	Protein			FT	Peptide	189..193	/label= Immunogenic_peptide	FT
XX				FT	Peptide	194..198	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	199..203	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX				FT	Peptide	204..208	/label= Immunogenic_peptide	FT
XX	Domain			FT	Peptide	204..208	/label= Immunogenic_peptide	FT

FT	Peptide	/note= "Specifically claimed in claim 15"
FT	Peptide	/label= Immunogenic_peptide
FT	Peptide	/note= "Specifically claimed in claim 15"
FT	Peptide	24...33
FT	Peptide	/note= Immunogenic_peptide
FT	Peptide	/note= "Specifically claimed in claim 15"
FT	Peptide	17...33
FT	Peptide	/label= Immunogenic_peptide
FT	Peptide	/note= "Specifically claimed in claim 15"
FT	Peptide	66...73
FT	Peptide	/label= Immunogenic_peptide
FT	Peptide	/note= "Specifically claimed in claim 15"
FT	Peptide	103...108
FT	Peptide	/label= Immunogenic_peptide
FT	Peptide	/note= "Specifically claimed in claim 15"
FT	Peptide	190...197
FT	Peptide	/label= Immunogenic_peptide
FT	Peptide	/note= "Specifically claimed in claim 15"
FT	Peptide	202...215
FT	Peptide	/label= Immunogenic_peptide
FT	Peptide	/note= "Specifically claimed in claim 15"
FT	Peptide	190...215
FT	Peptide	/label= Immunogenic_peptide
FT	Peptide	/note= "Specifically claimed in claim 15"
FT	Peptide	124...133
FT	Peptide	/label= Repeat_1
FT	Peptide	134...138
FT	Peptide	/label= Repeat_2
FT	Peptide	139...143
FT	Peptide	/label= Repeat_3
FT	Peptide	144...148
FT	Peptide	/label= Repeat_4
FT	Peptide	149...153
FT	Peptide	/label= Repeat_5
FT	Peptide	154...158
FT	Peptide	/label= Repeat_6
FT	Peptide	159...163
FT	Peptide	/label= Repeat_7
FT	Peptide	164...168
FT	Peptide	/label= Repeat_8
FT	Peptide	169...173
FT	Peptide	/label= Repeat_9
FT	Peptide	174...178
FT	Peptide	/label= Repeat_10
FT	Peptide	179...183
FT	Peptide	/label= Repeat_11
FT	Peptide	184...188
FT	Peptide	/label= Repeat_12
FT	Peptide	189...193
FT	Peptide	/label= Repeat_13
FT	Peptide	194...198
FT	Peptide	/label= Repeat_14
FT	Peptide	199...203
FT	Peptide	/label= Repeat_15
FT	Peptide	204...208
FT	Peptide	/label= Repeat_16
PN	US2002173027-A1.	
XX	21-NOV-2002.	
XX	03-AUG-2001; 2001US-0922469.	
XX	17-MAR-1999; 99US-124820P.	
XX	17-MAR-2000; 2000US-0527345.	
XX	(ADLER/) ADLER D A.	
XX	(SHEP/) SHEPPARD P O.	
XX	Adler DA, Sheppard PO,	
XX	WPI; 2003-328428/31.	

DR N-PSDB; ABX93594.

XX Novel isolated zsig63 polypeptide, member of the adhesin family, useful
PT for treating dental caries, periodontal disease, thrush,
PT gastrointestinal disease, urinary tract infections, vaginal infections,
PT skin infections

XX Claim 10; Page 27-28; 32pp; English.

XX The invention relates to an isolated zsig63 polypeptide comprising at
CC least 90% identity to an amino acid sequence which comprises domain 1 of
CC zsig63, domain 2, domain 3, mature zsig63 and full length zsig63. Also
CC included are the polynucleotide encoding zsig63, a zsig63 expression
CC vector, a cultured cell comprising the vector and expressing the protein,
CC a DNA encoding a fusion protein (comprising amino acids 1-15, 16-37,
CC 38-126, 127-219 or 16-219 of zsig63 and an additional protein), using a
CC zsig63 reporter gene construct to identify zsig63 agonists, and
CC producing an anti-zsig63 antibody using zsig63 immunogenic peptides,
CC zsig63 is useful for detecting in a test sample, the presence of
CC antagonist of zsig63 protein activity. Zsig63 has antimicrobial
CC activity and since exhibits high expression in salivary gland, can be
CC used for treating dental caries, periodontal disease, thrush, and
CC gastrointestinal disease, urinary tract infections, vaginal infections,
CC skin infections and other epithelial wounds. The polypeptides can be
CC used to establish normal microflora and protect against pathogenic
CC colonization and invasion. Zsig63 can also be used for providing
CC pro-inflammatory activity for treating chronic, tissue damage
CC particularly in areas having limited or damaged vascular system, e.g.
CC in diabetes, and for treating immunocompromised AIDS patients or in
CC individuals that have undergone chemotherapy, radiation treatment, for
CC treating lung infections e.g. in cystic fibrosis. Detection of zsig63
CC polypeptide at relatively high levels in the trachea may indicate that
CC such polypeptides may serve as a marker of lung dysfunction. Zsig63 is
CC also useful in diagnosing conditions associated with salivary gland or
CC lung dysfunction including salivary gland carcinoma, Pneumocystis carinii
CC infection, emphysema, chronic bronchitis, prostate dysfunctions such
CC as prostatic adenocarcinoma, aiding digestion, and as components of
CC defined cell culture media and may be used to replace serum that is
CC commonly used in culture. The DNA is useful in gene therapy applications
CC to increase or inhibit zsig63 activity, and for detecting abnormalities
CC on human chromosome 4 (e.g. 4q12-4q13, associated with dentinogenesis
CC imperfecta, and dentin dysplasia type II). Zsig63 is an adhesin family
CC member. The present sequence represents human zsig63.

XX Sequence 219 AA;

Query Match 100.0%; Score 1178; DB 24; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.5e-85;

Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLMACTVCVAFARRRPPIGDDNDGHPHPSLNIPIGIRNLPPPLYRRVNTVP 60
DB 1 MLLLMACTVCVAFARRRPPIGDDNDGHPHPSLNIPIGIRNLPPPLYRRVNTVP 60
QY 61 SYRGNTYTDGTPSYWMLTSGFPYVYHIRGFPLATQNLVPLPRGFPFVPSRFSFA 120
DB 61 SYRGNTYTDGTPSYWMLTSGFPYVYHIRGFPLATQNLVPLPRGFPFVPSRFSFA 120
QY 121 AAAPAPPIAAEPAAAAPLTATPVAAPAGAPVAAEPAAEPVGAEPAAEPVAAEPAA 180
DB 121 AAAPAPPIAAEPAAAAPLTATPVAAPAGAPVAAEPAAEPVGAEPAAEPVAAEPAA 180
QY 181 EAPVGEPAABEPSPAEPAATAKPAADPHPSLSLEQANQ 219
DB 181 EAPVGEPAABEPSPAEPAATAKPAADPHPSLSLEQANQ 219

RESULT 6
AAW30653
ID AAW30653 standard; Protein; 219 AA.
XX
AC AAW30653;

DT 12-APR-1999 (first entry)

XX Human secreted protein clone cp16 1 protein.
XX
XX Human; secreted protein; nutritional activity; cytokine; vaccine;
KW cell proliferation; differentiation; immune stimulation; suppression;
KW haematopoiesis regulation; tissue growth; activin; inhibitor; chemotactic;
KW chemokinetic; haemostatic; thrombolytic; anti-inflammatory; gene therapy;
KW tumour invasion suppression; tumour inhibition.

XX Homo sapiens.

XX MO9901466-A1.

XX 14-JAN-1999.

XX 01-JUL-1998; 98WO-US13813.

XX 27-OCT-1997; 97US-0958304.

XX 02-JUL-1997; 97US-0887195.

XX (GENY) GENETICS INST INC.

PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Racie LA;

PI Spaulding V, Treacy M;

DR WPI; 1999-105994/09.

DR N-PSDB; AAV80740.

PT New polynucleotides encoding secreted human proteins - are derived
PT from human foetal brain, adult testis, adult brain, foetal kidney,
PT adult salivary gland, or adult blood cDNA libraries, useful as, e.g.
PT potential vaccines

PS Claim 24; Page 71-72; 107pp; English.

XX The present sequence represents a human secreted protein from clone
CC cp16 1, deposited as ATCC 98482. Human secreted protein clone
CC polynucleotides and proteins are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals. Suggested
CC activities include nutritional activity, cytokine and cell
CC proliferation/differentiation activity, immune stimulating (e.g. as
CC vaccine) or suppressing activity, haematopoiesis regulating activity,
CC tissue growth activity, activin/inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC ligand activity, anti-inflammatory activity, cadherin/tumour invasion
CC suppressor activity, and tumour inhibition activity. The polynucleotides
CC are also stated to be useful for gene therapy.

XX Sequence 219 AA;

Query Match 99.5%; Score 1172; DB 20; Length 219;
Best Local Similarity 99.1%; Pred. No. 7.4e-85;

Matches 217; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLMACTVCVAFARRRPPIGDDNDGHPHPSLNIPIGIRNLPPPLYRRVNTVP 60
DB 1 MLLLMACTVCVAFARRRPPIGDDNDGHPHPSLNIPIGIRNLPPPLYRRVNTVP 60
QY 61 SYRGNTYTDGTPSYWMLTSGFPYVYHIRGFPLATQNLVPLPRGFPFVPSRFSFA 120
DB 61 SYRGNTYTDGTPSYWMLTSGFPYVYHIRGFPLATQNLVPLPRGFPFVPSRFSFA 120
QY 121 AAAPAPPIAAEPAAAAPLTATPVAAPAGAPVAAEPAAEPVGAEPAAEPVAAEPAA 180
DB 121 AAAPAPPIAAEPAAAAPLTATPVAAPAGAPVAAEPAAEPVGAEPAAEPVAAEPAA 180
QY 181 EAPVGEPAABEPSPAEPAATAKPAADPHPSLSLEQANQ 219
DB 181 EAPVGEPAABEPSPAEPAATAKPAADPHPSLSLEQANQ 219

RESULT 7
AAV19472
ID AAV19472 standard; Protein; 221 AA.
XX
AC AAV19472;
XX
DT 14-JUL-1999 (first entry)
XX
DE Amino acid sequence of a human secreted protein.
XX
KW Human secreted protein; cancer; tumour; neurodegenerative disorder;
KW developmental abnormality; fetal deficiency; blood disorder; leukemia;
KW immune system abnormality; autoimmune disease; hepatic disease;
KW renal disease; lymphoma; inflammation; allergy; ischemic shock;
KW Alzheimer's; cognitive disorder; schizophrenia; prostate disease;
KW obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease;
KW lung disease; chymus disease; digestive disorder; endocrine disorder;
KW infection; AIDS.
XX
OS Homo sapiens.
XX
FN WO9922243-A1.
XX
PD 06-MAY-1999.
XX
PE 23-OCT-1998; 98MO-US22376.
XX
PR 24-OCT-1997; 97US-0063387.
PR 24-OCT-1997; 97US-0062784.
PR 24-OCT-1997; 97US-0063088.
PR 24-OCT-1997; 97US-0063089.
PR 24-OCT-1997; 97US-0063090.
PR 24-OCT-1997; 97US-0063091.
PR 24-OCT-1997; 97US-0063092.
PR 24-OCT-1997; 97US-0063097.
PR 24-OCT-1997; 97US-0063098.
PR 24-OCT-1997; 97US-0063099.
PR 24-OCT-1997; 97US-0063100.
PR 24-OCT-1997; 97US-0063101.
PR 24-OCT-1997; 97US-0063109.
PR 24-OCT-1997; 97US-0063110.
PR 24-OCT-1997; 97US-0063111.
PR 24-OCT-1997; 97US-0063148.
PR 24-OCT-1997; 97US-0063386.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Brewer LA, Carter KC, Duan DR, Ebner R, Endress GA;
PI Peng P, Florence C, Florence KA, Greene JM, Janat P;
PI Kaye H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Wei Y, Young P;
XX
DR WPI; 1999-303069/25.
DR N-PSDB; AAX61352.
XX
XX New isolated human genes and the secreted polypeptides they encode
XX
PS Claim 11; Page 401-402; 546bp; English.
XX
XX The specification describes cDNA sequences (AAX61322-X61470) encoding
XX human secreted proteins (AAV19472-Y19590). The polynucleotides and their
XX corresponding secreted polypeptides are useful for preventing, treating
XX or ameliorating medical conditions, e.g. by protein or gene therapy.
XX Pathological conditions can also be diagnosed by determining the amount
XX of the polypeptides in a sample or by determining the presence of
XX mutations in the polynucleotides. Specific uses are described for each
XX of the polynucleotides, based on which tissues they are most highly
XX expressed in, and include developing products for the diagnosis or
XX treatment of cancer, tumours, neurodegenerative disorders, developmental
XX abnormalities and fetal deficiencies, blood disorders, leukemia,
XX diseases of the immune system, autoimmune diseases, hepatic and renal
XX disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's
XX and cognitive disorders, schizophrenia, prostate diseases, obesity,

CC disorders involving osteoclasts such as osteoporosis, arthritis or
CC malignancies, diseases of testes, lung or thymus, digestive/endocrine
CC disorders, infections and AIDS. The polypeptides are also useful for
CC identifying their binding partners.
XX
SQ Sequence 221 AA;
XX
Query Match 94.5%; Score 1113.5; DB 20; Length 221;
Best Local Similarity 95.0%; Pred. No. 3.1e-80;
Matches 209; Conservative 1; Mismatches 9; Indels 1; Gaps 1;
XX
QY 1 MKLLIMACIVCAFAKRRPFPIGDDNDGDHPLPSINIPYGRNLPPPLYRPVNTVP 60
DB 1 MKLLIMACIVCAFAKRRPFPIGDDNDGDHPLPSINIPYGRNLPPPLYRPVNTVP 60
QY 61 SYPGNTYDTGTPSYWMLTSPGPYVYHNGFPPLATOLNPPLPFRGPPFPSPRFSFA 120
DB 61 SYPGNTYDTGTPSYWMLTSPGPYVYHNGFPPLATOLNPPLPFRGPPFPSPRFSFA 120
QY 121 AAAPAPPIAEPAAAPLTAIPVAEPAPAGAPVAEPAAEP-VCAPPAEAPVAEPAA 179
DB 121 AAAPAPPIAEPAAAPLTAIPVAEPAPAGAPVAEPAXGCHLELEPAEAPVAEPAA 180
QY 180 AEPYGVPEPAEESPAPATATKPAAPPHPSPSLEQANQ 219
DB 181 AEPYGVPEPAEESPAPATATKPAAPPHPSPSLEQANQ 220
XX
RESULT 8
ABG21919
ID ABG21919 standard; Protein; 325 AA.
XX
AC ABG21919;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21910.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS86106.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX biologically -
XX
PS Claim 20; SEQ ID No 52278; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 325 AA;

Query Match 15.6%; Score 183.5; DB 22; Length 325;
 Best Local Similarity 34.3%; Pred. No. 9.5e-07;
 Matches 57; Conservative 8; Mismatches 78; Indels 23; Gaps 6;

QY 48 PPPLYYRPVNTVPSYSGNTTDTGLPSYWIITSGFPYVYHIRGFLATQINVPPLPR 107
 DB 45 PCPSRYCPQMTPEPP-----PAPPPLPPPP-----APASPPLPAPSP- 87
 QY 108 GPPFPVPSRFFSAAAPAPPIAAEPAAAPLTATVAAEPAAAGAVAAEPAAEPVGA 167
 DB 88 --PHAPPQALPPPPALPPPP--ASDPLLPALSPPLPAPPSAPAPPPAPPPAP 143
 QY 166 PAEAAPVAAEPAAEPVGAEPAAEPSPAPATAPAPAPPPSP 212
 DB 144 PPSPPAPPSAPSPAPLPAPAPSPSPAPPPAPPPAPPPAP 108

RESULT 9
 AAR40797

ID AAR40797 standard; Protein; 207 AA.

XX AC AAR40797;

XX DT 25-MAR-2003 (updated)
 XX DT 08-FEB-1994 (first entry)

XX DE Sequence of a 10-amino acid repetitive sequence of an antigenic
 XX DE peptide domain encoded by an insert of clone Tcd.

XX KW Epitope repeat; antigen; Chagas disease; TCD.

XX OS Trypanosoma cruzi.

XX FT Key Location/Qualifiers

XX FT MISC-difference 42
 XX FT /label= Degeneracy in repeat unit
 XX FT /note= "see also AAs 88,108,158,162"

XX PN WO9316199-A1.

XX PD 19-AUG-1993.

XX PF 11-FEB-1993; 93WO-US01231.

XX PR 14-FEB-1992; 92US-0836642.

XX PA (REBD/) REBD S G.

XX PI Reed SG;

XX DR WPI; 1993-272900/34.

XX DR N-PSDB; AAR40797.

XX PT Antigenic peptide domain of trypanosoma cruzi - used to diagnose
 XX PT or immunise against infection, and screen blood supplies

PS Disclosure; Figure 1; 16pp; English.

XX CC Clone Tcd encodes a 10-amino acid repetitive sequence present
 CC in 20.5 copies with minor degeneracies present in 5 posns.
 CC The predicted mol. wt. of recombinant unglycosylated Tcd antigen
 CC is 36.3kD. Analysis of DNA from seven geographically diverse
 CC T. cruzi isolates indicate that Tcd gene sequence was conserved.
 CC A 636 bp fragment of clone Tcd was used to probe DNA from several
 CC other protozoan parasites of humans. However, it did not hybridise
 CC to T. brucei, Leishmania chagasi, L. amazonensis, L. donovani
 CC or T. rangeli.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 207 AA;

Query Match 15.3%; Score 180; DB 14; Length 207;
 Best Local Similarity 38.4%; Pred. No. 1.1e-06;
 Matches 43; Conservative 11; Mismatches 46; Indels 12; Gaps 2;

QY 114 PSRFSAAAPAP-----PIAAEPAAAPLTATVAAEPAAAGAPVAAEPAAEPV 164
 DB 85 PAESKSAEPKPAEPKPAEPKPAEPKPAEPKPAEPKPAEPKPAEPKPAEPKPAEPK 144
 QY 165 GAEPAAEPVAAEPAAEPVGAEPAAEPSPAPATAPAPAP--HSPS 213
 DB 145 PAEPKSAEPKPAEPKSAEPKPAEPKPAEPKPAEPKPAEPKPAEPKPAEPKPA 196

RESULT 10
 AAM06915

ID AAM06915 standard; Protein; 207 AA.

XX AC AAM06915;

XX DT 02-APR-1997 (first entry)

XX DE T. cruzi Tcd antigen.

XX KW TCD; diagnosis; infection; antigen; Chagas disease.

XX OS Trypanosoma cruzi.

XX FT Key Location/Qualifiers

XX FT Region 6..20
 XX FT /label= Antigenic epitope
 XX FT /note= "Claim 3, page 38"

XX PN WO9629605-A2.

XX PD 26-SEP-1996.

XX PF 12-MAR-1996; 96WO-US03380.

XX PR 14-MAR-1995; 95US-0403379.

XX PA (COR1-) CORIXA CORP.

XX PI Reed SG;

XX DR WPI; 1996-485445/48.

XX DR N-PSDB; AAT46149.

XX PT Detecting and preventing T. cruzi infection - using polypeptide(s)
 XX PT or antibodies contg. or reactive with antigen epitope(s) of T. cruzi
 XX PT proteins

XX PS Disclosure; Fig 4; 59pp; English.

XX CC The Tcd antigen (AAM06915) of Trypanosoma cruzi includes an
 CC antigenic epitope (see also AAM06917) that can be utilised, pref.
 CC with other T. cruzi epitopes (see also AAM06914-19), in immunoassays
 CC for T. cruzi infection (Chagas' disease) and in vaccine compns.

SQ Sequence 207 AA; 15.2%; Score 179; DB 17; Length 207;
 Query Match 41.0%; Pred. No. 1.3e-06;
 Best Local Similarity 41.0%; Pred. No. 1.3e-06;
 Matches 43; Conservative 11; Mismatches 45; Indels 6; Gaps 2;

QY 114 PSRFFSAAPAPAPPIAAAPLTTPTVPAABPAAGAPVAAEPAAEPVGAEPAAEP 173
 DB 5 PAEPKSAEPKPAEPK-SAEPPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP 63
 QY 174 VAAEPAAEPVGVPEPAEESP-----AEPAATKPADEPHPS 213
 DB 64 KPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAES 108

RESULT 11
 AAY32850
 ID AAY32850 standard; Protein; 207 AA.
 AC AAY32850;
 XX
 DT 29-OCT-1999 (first entry)
 XX
 DE Tcd protein sequence.
 XX
 KM TCE repeat sequence; 35KD T. cruzi homologue; immune system disorder;
 KM eukaryotic ribosomal protein L19E; infection; detection; diagnosis;
 KM Chagas' disease; Leishmania infection.
 OS Trypanosoma cruzi.
 XX
 PN US5942403-A.
 XX
 PD 24-AUG-1999.
 XX
 PF 15-SEP-1997; 97US-0929414.
 XX
 PR 15-SEP-1997; 97US-0929414.
 PR 14-MAR-1995; 95US-0403379.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Houghton R, Reed SG, Skeiky YAM;
 DR WPI; 1999-517419/43.
 DR N-PSDB; AA210979.
 XX
 PT Trypanosoma cruzi antigenic homologue of eukaryotic ribosomal
 PT protein L19E, useful for screening assays to detect T. cruzi
 PT infection
 XX
 PS Example 3; Fig 4; 26pp; English.
 XX
 CC This sequence is the Trypanosoma cruzi Tcd protein. Antigenic fragments
 CC of Tcd and fragments of the 35 kD Trypanosoma cruzi homologue (Tcd) of
 CC the eukaryotic ribosomal protein L19E can be used in the method of the
 CC invention for detecting T. cruzi infection in a biological sample, which
 CC comprises: (a) contacting a biological sample with a first polypeptide
 CC comprising the sequence shown in AAY32839; (b) contacting the biological
 CC sample with a second polypeptide comprising the Tcd sequence shown in
 CC AAY32840 or AAY32841; and (c) detecting the presence of antibodies that
 CC bind to at least one of the polypeptides, indicating T. cruzi infection.
 CC The methods and polypeptides may be used to detect T. cruzi (which
 CC causes Chagas' disease and a variety of immune system disorders) and/or
 CC Leishmania infection in individuals and blood supplies. The compounds
 CC and methods may also be used to protect against T. cruzi infection. The
 CC N-terminal region of Tcd (residues 1-136) cross-reacts with
 CC anti-Leishmania antibodies and can be specifically excluded from the
 CC antigenic polypeptide to avoid such cross-reactivity.
 CC
 CC Sequence 207 AA;
 SQ Query Match 15.2%; Score 179; DB 20; Length 207;

Best Local Similarity 41.0%; Pred. No. 1.3e-06;
 Matches 43; Conservative 11; Mismatches 45; Indels 6; Gaps 2;

QY 114 PSRFFSAAPAPAPPIAAAPLTTPTVPAABPAAGAPVAAEPAAEPVGAEPAAEP 173
 DB 5 PAEPKSAEPKPAEPK-SAEPPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP 63
 QY 174 VAAEPAAEPVGVPEPAEESP-----AEPAATKPADEPHPS 213
 DB 64 KPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAES 108

RESULT 12
 AAY23305
 ID AAY23305 standard; Peptide; 207 AA.
 AC AAY23305;
 XX
 DT 31-AUG-1999 (first entry)
 XX
 DE Tcd antigen of Trypanosoma cruzi.
 XX
 KM Trypanosoma cruzi epitope; Trypanosoma cruzi infection; antigen;
 KM vaccine; Chagas' disease; Tcd antigen.
 XX
 OS Trypanosoma cruzi.
 XX
 PN W09931246-A1.
 XX
 PD 24-JUN-1999.
 XX
 PF 04-DEC-1998; 98WO-US25871.
 XX
 PR 18-DEC-1997; 97US-0993674.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Houghton RL, Lodes MJ, McNeill PD, Reed SG, Skeiky YAM;
 PI Smith JM;
 DR WPI; 1999-405035/34.
 XX
 PT New isolated Trypanosoma cruzi epitopes
 PT
 PS Disclosure; Page 83-84; 103pp; English.
 XX
 CC The specification describes new Trypanosoma cruzi epitopes. A method
 CC for detecting Trypanosoma cruzi infection in a biological sample
 CC comprises contacting the sample with a polypeptide comprising
 CC an epitope of a TC antigen, or a variant of the antigen that differs
 CC only in conservative substitutions and/or modifications and detecting
 CC the presence of antibodies that bind to the polypeptide in the sample,
 CC thereby detecting TC infection. The TC polypeptides can be used in
 CC vaccines for inducing protective immunity against Chagas' disease in
 CC a patient. The polypeptides and antibodies can also be used for detecting
 CC TC infection. The present sequence represent a Tcd antigen, which
 CC can be used in the above assay to improve sensitivity.
 CC
 CC Sequence 207 AA;
 SQ Query Match 15.2%; Score 179; DB 20; Length 207;
 Best Local Similarity 41.0%; Pred. No. 1.3e-06;
 Matches 43; Conservative 11; Mismatches 45; Indels 6; Gaps 2;

QY 114 PSRFFSAAPAPAPPIAAAPLTTPTVPAABPAAGAPVAAEPAAEPVGAEPAAEP 173
 DB 5 PAEPKSAEPKPAEPK-SAEPPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP 63
 QY 174 VAAEPAAEPVGVPEPAEESP-----AEPAATKPADEPHPS 213
 DB 64 KPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAES 108

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:10:34 ; Search time 35 Seconds
(without alignments)
1614.672 Million cell updates/sec

Title: US-09-923-236-2

Perfect score: 1178

Sequence: 1 MKLLLMACIVAVFARKRRF.....TAKPAAPHPHPSLEQANO 219

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database :

SPREMBL.23.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_tvirus:*
16: sp_bacterioplasmid:*
17: sp_archaeoplasmid:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	359.5	30.5	344	11 Q9D5J3	Q9D5J3 mus musculus
2	224.5	19.1	3179	12 Q9V2A4	Q9V2A4 human herpes
3	221	18.8	1100	2 Q9X6J9	Q9X6J9 klebsiella
4	214.5	18.2	1480	10 Q9LIR8	Q9LIR8 arabidopsis
5	214	18.2	3105	12 Q8UZ19	Q8UZ19 cercopithec
6	206.5	17.5	607	5 Q17112	Q17112 babesia bov
7	204.5	17.4	1070	2 Q8GFS8	Q8GFS8 citrobacter
8	199	16.9	391	5 Q26892	Q26892 trypanosoma
9	199	16.9	391	5 Q26893	Q26893 trypanosoma
10	194.5	16.5	722	5 Q26893	Q26893 trypanosoma
11	193	16.4	3534	12 Q39266	Q39266 equine herpes
12	189.5	16.0	1108	2 Q49542	Q49542 mycoplasma
13	188	16.0	839	16 Q9R574	Q9R574 deinococcus
14	186.5	15.8	395	16 Q9R574	Q9R574 deinococcus
15	186.5	15.8	616	11 Q99K31	Q99K31 mus musculus
16	183	15.5	744	10 Q65375	Q65375 arabidopsis

17	182	15.4	386	10 Q9FP05	Q9FP05 chlamydomon
18	181	15.4	928	12 Q9IMX9	Q9IMX9 cercopithec
19	181	15.4	2657	11 Q88493	Q88493 mus musculus
20	178.5	15.2	243	10 Q9X1V1	Q9X1V1 cucumis sat
21	177.5	15.1	598	16 Q8VKN7	Q8VKN7 mycobacteri
22	177	15.0	1315	10 Q9SPM0	Q9SPM0 zea mays (m
23	176.5	15.0	3084	12 Q8UZ11	Q8UZ11 pseudorabie
24	176	14.9	959	16 Q8XRH0	Q8XRH0 ralsconia s
25	175	14.9	857	16 Q98IK2	Q98IK2 rhizobium l
26	174.5	14.8	979	4 Q8N4Z1	Q8N4Z1 homo sapien
27	174.5	14.8	1611	3 Q42854	Q42854 schizosach
28	174	14.8	236	10 Q9LM00	Q9LM00 pinus taeda
29	174	14.8	503	12 Q39779	Q39779 equine herp
30	173.5	14.7	786	10 Q48809	Q48809 arabidopsis
31	173	14.7	1174	4 Q94854	Q94854 homo sapien
32	171.5	14.6	438	16 Q9A2M5	Q9A2M5 caulobacter
33	171.5	14.6	616	4 Q9H6K5	Q9H6K5 homo sapien
34	171	14.5	544	12 Q89392	Q89392 parametium
35	171	14.5	1340	16 Q9LJH8	Q9LJH8 streptomyce
36	170.5	14.5	601	5 Q17113	Q17113 babesia bov
37	170	14.4	295	2 Q32850	Q32850 mycobacteri
38	170	14.4	428	10 Q941Q8	Q941Q8 zea mays (m
39	170	14.4	548	16 Q06404	Q06404 mycobacteri
40	170	14.4	842	5 Q9VG59	Q9VG59 drosophila
41	170	14.4	864	5 Q95U45	Q95U45 drosophila
42	170	14.4	864	5 Q9VGC8	Q9VGC8 drosophila
43	170	14.4	1354	11 Q9EPW8	Q9EPW8 mus musculu
44	169.5	14.4	802	11 P70433	P70433 mus musculu
45	169	14.3	312	5 Q9UAN1	Q9UAN1 drosophila

ALIGNMENTS

RESULT 1	ID	Q9D5J3	PRELIMINARY	PRT	344 AA.
AC	Q9D5J3	01-JUN-2001 (TRENBLrel. 17, Created)			
DT	01-JUN-2001 (TRENBLrel. 17, Last sequence update)				
DT	01-JUN-2001 (TRENBLrel. 17, Last annotation update)				
DE	4930432K09R1K protein.				
GN	4930432K09R1K.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=Testis;				
RX	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,				
RA	Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,				
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Guerinich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima J., Mazzaletti U., Mombert P.,				
RA	Norione P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmig L.,				
RA	Wynshaw-Borje A., Yoshida K., Haegawa Y., Kawaji H., Kohlsuki S.,				
RA	Hayashizaki Y.,				
RT	"Functional annotation of a full-length mouse cDNA collection."				
RL	Nature 409:685-690(2001).				
DR	EMBL; AK015291; BAB29782.1; -				
DR	MGI; MGI:1921029; 4930432K09R1K.				

DR InterPro: IPR002965; P rich extensn.

DR PRINTS; PRO1217; PRICHEXTENS. 344 AA; 36294 MW; D6SEAD71CE5802AC CRC64;

Query Match 30.5%; Score 359.5; DB 11; Length 344;
Best Local Similarity 32.0%; Pred. No. 8.8e-18;
Matches 106; Conservative 25; Mismatches 87; Indels 113; Gaps 11;

```

QY 1 MKLLMAGTCAVAFARRRPPFAGEDN--DDGHPLHPSLNIPYGI--RNLPP---PPLY 53
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MKLLMAGTCAVAFARRRPPFAGEDN--DDGHPLHPSLNIPYGI--RNLPP---PPLY 59
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 54 RPVNTVPSYGNVYTGTLPSYPMILTSQGFVYHIRGFLATOLNVPL----- 104
   :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 60 QCGNTTKPGSELEKGLSLYPWIAIPSKLRYVQSLVNPADAPLNGPVAFLPPKAF 119
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 105 PRGPFVPPSRFP-----SAAAPAAP----- 127
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 120 PRPYPFVIPPKISVSPVREPVAVPAMPAGEGLVPPVPDKILGLPQAVKGTVP 179
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 128 -----DIAEPA-----AAPLTATVAAEPAAGAVAAE 157
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 180 PSGKPLADEPAPSPGAPAPVQCGAPAPAHQLEAAVPPASRLMADEPAVPLSVGAQ 239
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 158 PAEAPVQ-----AEPAAAPVAAEPAAPV-----GVEPAEESP 195
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 240 SLAESPAIGSPENKPTSGSPATQSPALPAAGLAVBAKLPAAESAGRPPEALMASOS 299
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 196 -----AEPATKPAAPPPSPSLEQANQ 219
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 300 VVGKLTAEPTAKPOVLEVEAKSASQEAQ 330

```

RESULT 2

Q8V2A4 PRELIMINARY; PRT; 3179 AA.

```

AC Q8V2A4; PRELIMINARY; PRT; 3179 AA.
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE BPLF.
OS Human herpesvirus 4 (Epstein-Barr virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10376;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M-ABA;
RT "Schmaus S., Wolf H., Schwarzmann F.;
   "The open reading frame BPLF of Epstein-Barr virus is expressed in
   late viral replication."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF448220; AAL40834.1; -.
DR InterPro: IPR006928; Herpes teg N.
DR InterPro: IPR002965; P rich extensn.
DR Pfam; PF04843; Herpes_teg_N; 1.
DR PRINTS; PRO1217; PRICHEXTENS.
SQ SEQUENCE 3179 AA; 340397 MW; AB0ASD9088A945B1 CRC64;

```

```

Query Match 19.1%; Score 224.5; DB 12; Length 3179;
Best Local Similarity 31.1%; Pred. No. 2e-07;
Matches 68; Conservative 23; Mismatches 61; Indels 67; Gaps 9;
QY 24 GEDNDGDHPLHPSLNIPYGI--RNLPP---PPLY 83
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 278 GEAAGTGGADSSPWLPRFRKIXPNL-----PSRFTS---DSFAARYSP 325
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 84 PPVYHIRGFLATOLNVPLPRGSPVPSRFFSAAPAAAPVAAEPAAPVAT 142
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 326 -----AKTNSPPSSP-----ASAPASAPASAPASAPASAPAS 359
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 143 PVAAEPAAGPVAEPAEAPVGAEPAAEAPVAAEPAEAPVVEEPAEAP 193
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

DB 360 PASAAPASAPASAPASAPASAPASAPASAPASAPASAPASAPASAPASAP 419

QY 194 -SPAEPATAKP-----AAPF-----PRPS 213
420 HTPGVAPSTPSPRASGAGAPOTPRKRKGLGKDSPKKPT 458

RESULT 3

Q9X6J9 PRELIMINARY; PRT; 1100 AA.

```

ID Q9X6J9; PRELIMINARY; PRT; 1100 AA.
AC Q9X6J9;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE DNA primase.
GN F1.
OS Klebsiella oxytoca.
OC Plasmid pACM1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=571;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20334746; PubMed=1087523;
RA Preston K.E., Radomski C.C.A., Venezia R.A.;
RT "Nucleotide sequence of a 7-kb fragment of pACM1 encoding an IncM DNA
   primase and other putative proteins associated with conjugation."
RL Plasmid 44:12-23(2000).
DR EMBL; AF139719; AAD31802.1; -.
KW Plasmid.
SQ SEQUENCE 1100 AA; 120370 MW; 36EAD2BD33FCB732 CRC64;

```

Query Match 18.8%; Score 221; DB 2; Length 1100;
Best Local Similarity 31.4%; Pred. No. 1.3e-07;
Matches 76; Conservative 19; Mismatches 85; Indels 62; Gaps 9;

```

QY 9 IYCVAFARRRRP-----FIGEDN--NDP-----GHPLHPSLNIPYGIN 46
   :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 296 MEVAGRLKDTFNSNTHYFLADNDYKDENVGLKATAEATAGHVLVPASNN----- 352
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 47 LPPLYPVPVNTVPSYGNVYTGTLPSYPMILTSQGFVYHIRGFLATOLNVPLP 105
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 353 -----KEGLTDYNDLVSEGLQVQLQVEG--AINQMRVDTM 388
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 106 PRGPFVPPSRFFSA-AAAPAP-----PLAEPAAAPVATPVAAEPAAGPVAEPA 160.
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 389 PTDNPNITDVNHSSTDSAAVAAPBEKAAAPASTPAAAEVEEAPVASAPAAEVAAPVA 448
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 161 EAPVGAE-----PAEAPVAAEPAEAPVGAEPAAEESPAPAPATKAPAPPPSPS 215
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 449 SAPAAAEVETAPVASTPEAAEPVAAEPVAAEPAEPAEPAEPAEPAEPAEPAE 508
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 216 QA 217
DB 509 SA 510

```

RESULT 4

Q9LIE8 PRELIMINARY; PRT; 1480 AA.

```

ID Q9LIE8; PRELIMINARY; PRT; 1480 AA.
AC Q9LIE8;
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Similarity to cell wall-plasma membrane linker protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;

```

RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the regions of 4,251,635 bp covered by ninety pl,
TAC and BAC clones.";
RL DNA Ref. 7:217-221(2000).
DR EMBL; AP001306; BAB03062.1; -
DR HSSP; P24337; IHP.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR002965; P_rich_extensn.
DR Pfam; PF00234; tryp_alpha_amy1.1.
DR PRINTS; P01217; PRICEXTENS.
SQ SEQUENCE 1480 AA; 147154 MW; D1AC0C79F155E732 CRC64;
Query Match 18.2%; Score 214.5; DB 10; Length 1480;
Best Local Similarity 32.6%; Pred. NO. 4.8e-07; Indels 47; Gaps 9;
Matches 63; Conservative 13; Mismatches 70;
QY 33 PLHPSLNIPYGRNLP---PPLYRPNVTVPSTPGNTYDTGLPSYPMILSPGFPYVYH 89
DB 665 PAKPVPVATP-----PIATPPVAKPVPATPP-----TATPP-----IATP----- 698
QY 90 IRGFPLATQLVNPPV--PPRGFPV-----PSGRFSAAPAAAP--PPIAAEPAAAPL 139
DB 699 ----PIAT-----PVPVTPPTATPPVATPPPIAKPPTTIPPTATPPVAMPPPIATPPAKKPI 750
QY 140 TATPPAAPPAAGAPPAAPPAAPPAAPPAAPPAAPPAAPPAAPPAAPPAAPPAAPPA 199
DB 751 ATPPIATPPVAKPVPVTPPTATPPPIATPPPIAKSPVATPPATPPVATPPVATPPPT 810
QY 200 TAKPAAPPPHPSP 212
DB 811 TAPPTATPPVAKP 823
RESULT 5
Q8UZ19 PRELIMINARY; PRT; 3105 AA.
ID Q8UZ19
AC Q8UZ19
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE BP1F1.
OS Cercopithecine herpesvirus 15.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=104228;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=97048062; PubMed=8892903;
RA Franken M., Devergne O., Rosenzweig M., Annis B., Kieff E., Wang F.;
RT "Comparative analysis identifies conserved tumor necrosis factor
receptor-associated factor 3 binding sites in the human and simian
Epstein-Barr virus oncogene LMP1.";
RL J. Virol. 70:7819-7826(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=99412410; PubMed=10482645;
RA Rivallier P., Quink C., Wang F.;
RT "Strong selective pressure for evolution of an Epstein-Barr virus
LMP2B homologue in the rheus lymphocryptovirus.";
RL J. Virol. 73:8867-8872(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=20304984; PubMed=10846073;

RA Jiang H., Cho Y.-G., Wang F.;
RT "Structural, functional, and genetic comparisons of Epstein-Barr virus
nuclear antigen 3A, 3B, and 3C homologues encoded by the rheus
lymphocryptovirus.";
RL J. Virol. 74:5921-5932(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=20440633; PubMed=10970361;
RA Rao P., Jiang H., Wang F.;
RT "Cloning of the rheus lymphocryptovirus viral capsid antigen and
Epstein-Barr virus-encoded small RNA homologues and use in diagnosis
of acute and persistent infections.";
RL J. Clin. Microbiol. 38:3219-3225(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=21602573; PubMed=11739708;
RA Rivallier P., Jiang H., Cho Y.-G., Quink C., Wang F.;
RT "Complete Nucleotide Sequence of the Rheus Lymphocryptovirus: Genetic
Validation for an Epstein-Barr Virus Animal Model.";
RL J. Virol. 76:421-426(2002).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RA Moghaddam A., Koch J., Annis B., Wang F.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RA Moghaddam A., Annis B., Wang F.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RA Rivallier P., Quink C., Wang F.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RA Jiang H., Wang F.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [10]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RA Rao P.V., Jiang H., Wang F.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RA Rivallier P., Jiang H., Cho Y.-G., Quink C., Wang F.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY037858; AAK95420.1; -
DR InterPro; IPR005613; AIP3.
DR InterPro; IPR006928; Herpes teg N.
DR InterPro; IPR002965; P_rich_extensn.
DR Pfam; PF03915; AIP3.1.
DR Pfam; PF04843; Herpes teg N.1.
DR PRINTS; P01217; PRICEXTENS.
SQ SEQUENCE 3105 AA; 332890 MW; 7C6AED7D99FC7C3 CRC64;
Query Match 18.2%; Score 214; DB 12; Length 3105;
Best Local Similarity 29.6%; Pred. NO. 1.1e-06;
Matches 61; Conservative 18; Mismatches 105; Indels 22; Gaps 7;
QY 18 RRPFGIDNDGDGHLPSLNIPYGRNLPPLYV--RPVNTVPSTPGNTYDTGLPSY 75
DB 390 RRLPMSSTTDEDQD--PPRRPSPVLTPLAPVTVVQGPAPVPTQPSAAPPRA 447
QY 76 PWILTSPGFPYVYHIR---GFPLATQLVNPPVLPGRGFPVPPSRRFSAAPAAAP--PI 129
DB 448 P-----SPVIRPHSPTRKPPDPPTTSQLPPTQPPPISTPLVLP-----ISTPLVPRPT 497

QY 130 AAEPAAPLTAATPVAAEPAAAGAPVAAEPAAAPVGAEPAAAPVAAEPAAAPVGEPA 189
 DB 498 POPPAATPTPOPPOAATPTPOPPOAATPTPOPPOAATPTPOPPOAATPTPOPPOAATPT 557
 QY 190 AEEPPAE--PATAPKAPAEPPHSPS 213
 DB 558 POPPOATSHAPQPPRAASAPAPQPTPT 583

RESULT 6

ID 017112 PRELIMINARY; PRT; 607 AA.
 AC 017112:
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE 80 kDa protein.
 OS Babesia bovis.
 OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Babesiidae; Babesia.
 OX NCBI_TaxID=5865;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93341567; PubMed=8341317;
 RA Dalrymple B.P., Peters J.M., Goodger B.V., Bushell G.R.,
 RA Waltsbuhl D.J., Wright I.G.,
 RT "Cloning and characterisation of cDNA clones encoding two Babesia
 RT bovis proteins with homologous amino- and carboxy-terminal domains."
 RT Mol. Biochem. Parasitol. 59:181-189 (1993).
 RL EMBL; M93126; AAA02753.1; -
 SR SEQUENCE 607 AA; 67129 MW; 2E06ECTDF843D732 CRC64;

Query Match 17.5%; Score 206.5; DB 5; Length 607;
 Best Local Similarity 46.5%; Pred. No. 7.3e-07;
 Matches 46; Conservative 7; Mismatches 35; Indels 11; Gaps 2;
 QY 128 PTAEPAAAPLTATPVAAEPAAAGAPVAAEPAAAPVGAEPAAAPVAAEPAAAPVGE 187
 DB 329 PVEEPVAAEPPIVEEPIVEEPIVEEPIVEEPIVEEPIVEEPIVEEPIVEEPIVEE 388
 QY 188 PAEEP-----SPAEPATAKPA---APEPPHSPSLR 215
 DB 389 PVAEPPVAEEPCETPAETPAETPAETPAETPAETPAETPAETPAETPAETPAETPAE 427

RESULT 7

ID 08GFS8 PRELIMINARY; PRT; 1070 AA.
 AC 08GFS8:
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE DNA primase.
 GN PRI.
 OS Citrobacter freundii.
 OC Plasmid pCTX-M3.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Citrobacter.
 OX NCBI_TaxID=546;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Goldbiewski M., Zienkiewicz M., Adamczyk M., Kern-Zdanowicz I.,
 RA Ceglowski F.,
 RT "Complete nucleotide sequence of highly transmissible plasmid pCTX-
 RT M3."
 RT Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF550415; AAN87680.1; -
 KW Plasmid.
 SQ SEQUENCE 1070 AA; 117583 MW; F16IDCB3137E6500 CRC64;

Query Match 17.4%; Score 204.5; DB 2; Length 1070;
 Best Local Similarity 30.9%; Pred. No. 1.7e-06;
 Matches 69; Conservative 16; Mismatches 81; Indels 57; Gaps 7;

QY 9 IVCVAFARKRRP-----FIGBD--NDD-----GHPHPSLINTPIYGIN 46
 DB 298 MEEVAGRLKDTFFPNSHTYFLDNDIYKDNENGLEKATTAELTACHVLPVPSN----- 352
 QY 47 LPPPLYRPPVNTVPSYGNITYDTGLPSYPILTSPPPYV-YHIRGFPLATQLVNPLP 105
 DB 353 -----KEGLTDYNDLHVSEGLEQVRLQVEG--AINQMRVDTM 388
 QY 106 PRGFPPVPPSRFFSAAAPPAAPPIAAEPAAAPLTATPVAAEPAAAGAPVAAEPAAAPV 165
 DB 389 PTDNENITDVNHSSTDSA-----AAVAPEKAPVASTPAAPVAAEPVAAAPVAAAPAAE 443
 QY 166 AEPAAEPVAAEPAAAPVGEPAEPAAEPSPAPATAPKAPAP 208
 DB 444 AAPVASAPAAEPVETAPVASTPEAAEPVEAPVAPVASAPAAAP 486

RESULT 8

ID 096579 PRELIMINARY; PRT; 796 AA.
 AC 096579:
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Surface antigen PHG87#5 (Fragment).
 OS Trypanosoma cruzi.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5693;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tulahuén;
 RA Garcia G.A., Bonetemp E., Bua J., Ruiz A.M.;
 RT "Molecular characterization of a Trypanosoma cruzi clone recognized by
 RT an anti-Schistosoma japonicum glutathione-S-transferase serum."
 RT Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF091620; AAC98688.2; -
 FT NON TER 1
 SQ SEQUENCE 796 AA; 83983 MW; DD4DCC06622533D8 CRC64;

Query Match 17.3%; Score 204; DB 5; Length 796;
 Best Local Similarity 30.5%; Pred. No. 1.4e-06;
 Matches 62; Conservative 15; Mismatches 62; Indels 64; Gaps 7;

QY 47 LPPPLYRPPV-----NTVPSYP-----GNTYTDGLPSYPILTS 82
 DB 411 IPPP-ERRKVPAAATSSSVBAPNERVTNTQPTVPSPATAGPQGT----- 456
 QY 83 GPPVYHIRGFPLATQLVNPLPPRGPFVP-----PSRFFSAAAPPAAPPIAEP 133
 DB 457 -----OTTVAASVPSGAGSKPAEPKSAEPKAPKAPKAPKAPKAPKAPKAPKAPKAP 502
 QY 134 AAAAPLTATPVAAEPAAAGAPVAAEPAAAPVGAEPAAAPVAAEPAAAPVGEPAEP 193
 DB 503 KPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP 562
 QY 194 SPAEPATAKPAEP---HPSPS 213
 DB 563 KPAEPKSAEPKPAEPKSAEPKPT 585

RESULT 9

ID 026892 PRELIMINARY; PRT; 391 AA.
 AC 026892:
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Surface antigen (Fragment).
 OS Trypanosoma cruzi.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5693;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC STRAIN=CAI;
RX MEDLINE=92389985; PubMed=1518528;
RA "Buochiazzo A., Campetella O.E.;"
RT "Sequence of the gene for a Trypanosoma cruzi protein antigenic during
   the chronic phase of human Chagas disease."
RL Mol. Biochem. Parasitol. 54:125-128 (1992).
DR EMBL; M92046; AAB97873.1; -.
FT NON TER
SQ SEQUENCE 391 AA; 39459 MW; F234844B1F865BC7 CRC64;

Query Match
Best Local Similarity 43.8%; Pred. No. 1.6e-06;
Matches 42; Conservative 14; Mismatches 40; Indels 0; Gaps 0;

QY 113 PPSRFSAAAPAPPIAEPAAAPLTPTVAEPAAAGAPVAEPAAEPVGAEPAAEA 172
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 PKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAE 64

QY 173 PVAEPAAEPVGEPAEPAPPAEPATKPAAP 208
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 PKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAE 100

RESULT 10
Q26893 PRELIMINARY; PRT; 722 AA.
AC Q26893;
AT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DR 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Surface antigen (Fragment).
OS Trypanosoma cruzi.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCB1_TaxID=5693;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CAI;
RX MEDLINE=92389985; PubMed=1518528;
RA "Buochiazzo A., Campetella O.E.;"
RT "Sequence of the gene for a Trypanosoma cruzi protein antigenic during
   the chronic phase of human Chagas disease."
RL Mol. Biochem. Parasitol. 54:125-128 (1992).
DR EMBL; M92047; AAB97874.1; -.
FT NON TER
SQ SEQUENCE 722 AA; 74874 MW; C572189C6DB4B3E CRC64;

Query Match
Best Local Similarity 29.8%; Pred. No. 5.9e-06;
Matches 59; Conservative 26; Mismatches 74; Indels 39; Gaps 6;

QY 20 PFFIGEDNDGHPHPSINIPYGRNLPPLYRPPVNTVP-----SYPGNTYTD 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 FYFGAYDQSLSRKIHATLANVF-----LYNRLNDALIGALNANKVSLP---FTEK 237

QY 71 GLPSPFWLTSRGFYVYHIRFPLATQUNVPLPRRGPFPPPSRFSAAAPAPPIA 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 -KPAASATATSPSVSHV-----TTPVATE-----PKSAEPKSAVPKSAEPKS 278

QY 131 AEPAAAPLTPTVAEPAAAGAPVAEPAAEPVGAEPAAEPVAAEPAAEPVGEPA 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 279 AVPKSAEPKSAEPKSAEPKSAVPKSAEPKSAVPKSAEPKSAEPKSAEPKSAEPKS 338

QY 191 EEPSPAEPATAKPAAP 208
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 339 AEPKSAEPKSAVPKSAEP 356

RESULT 11
Q39266 PRELIMINARY; PRT; 3534 AA.
AC Q39266;
AT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

```

```

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Counterpart of HSV-1 gene UL36 and VZV gene 22.
GN 24.
OS Equine herpesvirus 4.
OC Alphaherpesvirinae; no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirinae.
OX NCB1_TaxID=10331;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NS80567;
RX MEDLINE=96264497; PubMed=9603335;
RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-4."
RL J. Gen. Virol. 79:1197-1203 (1998).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NS80567;
RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF030027; AAC59539.1; -.
DR InterPro; IPR006828; Herpes_teg_N.
DR InterPro; IPR005210; Herpes_UL36.
DR InterPro; IPR002965; P_rich_exten.
DR Pfam; PF04843; Herpes_teg_N; 1.
DR Pfam; PF03586; Herpes_UL36; 1.
DR PRINTS; PRO1217; PRICHEXTENSN.
DR SEQUENCE 3534 AA; 379956 MW; AF5164843CA1BFC6 CRC64;

Query Match
Best Local Similarity 37.4%; Pred. No. 3.5e-05;
Matches 70; Conservative 22; Mismatches 67; Indels 28; Gaps 12;

QY 36 PSLNIPYGRNLPPLYRPPVNTVSPYNGNTYTDGLSYPIILTSRGFPYVYHIRGFL 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2697 PGYSIP--MDGLPPDPNEALITAPSKP-----AAAPAPSKPAAPAPSKP-----AAAP 2745

QY 96 ATQUNVPLPRRGPFVPPSRFSAAAPAPPIAEP-----AAAPLTATPVAA-----E 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2746 PSKPAAPAPSKPAAPAPSKP---PAAAPAPSKPAAPAPSKPAAPAPSKPAAPAPSK 2802

QY 148 PAAGAPVAEPAAEPVGAEPAAEPVAAEPAAEPVGAEPAAEPVGAEPAAEPVGA 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2803 PAA-APAPSKPAA-APAPSKPAA-APAPSKPAA-APAPSKPAA-APAPSKPAA-APAPSKP 2857

QY 207 EPHSPSPS 213
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2858 AAAPAPSPS 2864

RESULT 12
Q49542 PRELIMINARY; PRT; 1108 AA.
AC Q49542;
AT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DR 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Ciliary adhesion protein P97 (Cilium adhesion P97).
GN P97.
OS Mycoplasma hyopneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCB1_TaxID=2099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=232;
RX MEDLINE=97175562; PubMed=9023217;
RA Hsu T., Artushin S., Minion F.C.;
RT "Cloning and functional analysis of the P97 swine cilia adhesion gene
   of Mycoplasma hyopneumoniae."
RL J. Bacteriol. 179:1317-1323 (1997).
RN [2]
RP SEQUENCE OF 1089-1108 FROM N.A.
RC STRAIN=232A;
RA Hsu T., Minion F.C.;

```

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; U50901; AAB47806.1; -
DR EMBL; AF012905; AAC32526.1; -
DR InterPro; IPR000719; Prot_kinase.
DR PROSITE; PS00107; PROTEIN_KINASE_ARP; 1.
SQ SEQUENCE 1108 AA; 124903 MW; 239CF62D61E11FE7 CRC64;

Query Match 16.0%; Score 188.5; DB 2; Length 1108;
Best Local Similarity 51.2%; Pred. No. 2.4e-05;
Matches 42; Conservative 8; Mismatches 27; Indels 5; Gaps 1;

QY 124 PAAPPIAEPAAAPLTATPTVAEPAAAPVGAEPAAEPVGAEPAAEPVGAEPAAEP 183
DB 813 PAAPPAAPVAAKPEPTTTPVAAKPE-----AAKPAAPVAAKPEAAAPVAAKPEAAAP 667
QY 184 VGEPAAEPPSPAEPAATAPAA 205
DB 868 VAAKPEAAKPVAAKPEAAKPVAA 889

RESULT 13

Q9RX57 PRELIMINARY; PRT; 839 AA.
AC Q9RX57;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein DR0458.
GN DR0458.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;

RA White O., Eissen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
Vamathevan J.J., Lam P., McDonald L., Uteerback T., Zalewski C.,
Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
Fraser C.M.;
RA "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RT Science 286:1571-1577(1999).
DR EMBL; AE001904; AAF10038.1; -
DR TIGR; DR0458; -
DR InterPro; IPR002965; P rich_extensn.
DR PRINTS; PR01217; PRICHTXTNSN.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 839 AA; 79759 MW; 3B6C2CD0CFEFADE CRC64;

Query Match 16.0%; Score 188; DB 16; Length 839;
Best Local Similarity 33.0%; Pred. No. 2e-05;
Matches 64; Conservative 14; Mismatches 92; Indels 24; Gaps 9;

QY 27 DNDGDGHLPSLNIPIGIRNLPPPLTYRPVNTVPSYPGNTYTDGSPSYWILTSP-GFP 85
DB 140 DPTPAAPLKP-----VQDTPPVTPPKVTPPEPVTPKAPTPPEVTLQPPVAGTPVAKP 193
QY 86 YYTHIRGFPPLAT-QLVNPPLPFGFPFVPPSRFSAAPPAAPPIAAEPAAAPLTATPV 144
DB 194 PV-----PAFTSQTPPTPVQAPARTPTPPQ--AARPTNAPAQTPAPATQAPAAQTPT 244
QY 145 AAEBAAGAPVAAEPAAEPVGAEPAAEPVGAEPAAEPVGAEPAAEPVGAEPAAEPVGAEP 203
DB 245 AQAAPATQTPPTPAAPAAQAPGAP---SPAAPAAQANAPAGSVVPEATVPESTPA--P 299
QY 204 AAEPPHPSPLDQA 217
DB 300 SAQT-P-PTPTRETA 312

RESULT 14

Q9RV74 PRELIMINARY; PRT; 395 AA.
AC Q9RV74;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein DR0938.
GN DR0938.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;

RA White O., Eissen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
Vamathevan J.J., Lam P., McDonald L., Uteerback T., Zalewski C.,
Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
Fraser C.M.;
RA "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RT Science 286:1571-1577(1999).
DR EMBL; AE001946; AAF10516.1; -
DR TIGR; DR0938; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 395 AA; 40003 MW; 75FD963717DDAFAFO CRC64;

Query Match 15.8%; Score 186.5; DB 16; Length 395;
Best Local Similarity 35.0%; Pred. No. 1.2e-05;
Matches 57; Conservative 11; Mismatches 64; Indels 31; Gaps 9;

QY 60 PSYV-GNTYTDGLSYPIILTSPPGPPYVYHNGFPLATQLVNPPLPFGFPFVPPSRF 118
DB 53 PAEPVGAPOPTTSIPVSPAAVADSPQT---AAPAQQTAGKIPV--PAAP----- 102
QY 119 SAAAPAAP-----IAEPAAAPLTATPTVAEPAAAPVGAEPAAEPVGAEPAAEPVGAEP 170
DB 103 ----APKTPPVLPPEPRKVPAPPTPRPVETTTAAPTQGPVQAAPATQTPPTQTPAA 158
QY 171 EAPVAAEPAAEPVGEVPA--EPPSPA--EPAAT-AKPAAPP 208
DB 159 QAPATQPPATPAP---EPAPAPAPAPATPPEBATPTEPAAPAP 198

RESULT 15

Q99K31 PRELIMINARY; PRT; 616 AA.
AC Q99K31;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Similar to procollagen, type VI, alpha 3 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RA "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RT Science 286:1571-1577(1999).
DR EMBL; BC005491; AAH05491.1; -
DR HSSP; P12111; 1KXT.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR002223; Kunitz_BPTI.
DR InterPro; IPR002965; P rich_extensn.
DR InterPro; IPR002035; VWF_A.

DR Pfam; PF00041; fn3; 1.
 DR Pfam; PF00014; Kunitz_BPTI; 1.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PR00759; BASICPTASE.
 DR PRINTS; PR01217; PRICHEXTENS.
 DR PRODOM; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00060; FN3; 1.
 DR SMART; SM00131; KU; 1.
 DR SMART; SM00327; vwa; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS50279; BPTI_KUNITZ_2; 1.
 DR PROSITE; PS50234; vwa; 1.
 KW Collagen; Protease inhibitor; Serine protease inhibitor.
 FT NON TER
 SQ SEQUENCE 616 AA; 66408 MW; FBA6CAC92B853BB3 CRC64;

Query Match 15.8%; Score 186.5; DB 11; Length 616;
 Best Local Similarity 32.4%; Pred. No. 1.8e-05;

Matches 67; Conservative 25; Mismatches 72; Indels 43; Gaps 12;

OY	21	PTIGEDNDGDHPLHPSLNIPYGINLPPLTYRPNVTVPSTYNGNTYTDGLPSYFWILT	80
DB	253	PVLAKPDPAKPAQAPAPAKPASAKLVDP---QPVHQPAPAQTAQSVAPAPAKP----	303
OY	81	SPGFYVYHNGFPLATQLVNVPPLPRGSPFYV---PSRFSSAAAPAPPIAAEPAA	136
DB	304	APPOPA-----AKVPAPK--FAVPAQAPAPQAPAAKVPAPKPVAPQAPAA	348
OY	137	APLTATPV-----AAEPA-AGAPVAAEPAA--EAPVGAEPAAAEAPVAAEPAAEPV--G	185
DB	349	QPMAPQPLVLTSAATKPAASANKPVAAKPVATNTATATARPALAAKPAAPAAKPAATRPPLAA	408
OY	186	VEPAAEPP-----SPAEPATAKPAPEP	208
DB	409	VRPVATKEAPEARQAPAPATKPAATKPAATKP	435

Search completed: January 15, 2004, 15:14:11
 Job time : 37 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:06:54 : Search time 18 seconds

(without alignments)
572.158 Million cell updates/sec

Title: US-09-923-236-2

Perfect score: 1178

Sequence: 1 MKLLIMACIVCVAFARKRRF.....TAKPAPRPHPSLEQANQ 219

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	180	15.3	555	1	GP1_CHLRE
2	175	14.9	1054	1	IF2_STRAU
3	175	14.9	3149	1	TBGT_BBV
4	174	14.8	518	1	TPW4_DROME
5	170	14.4	3176	1	CA36_HUMAN
6	169.5	14.3	802	1	ENAH_MOUSE
7	169	14.3	299	1	RL22_DROME
8	169	14.3	865	1	CPN_DROME
9	167.5	14.2	450	1	CY1_PARDE
10	166	14.1	810	1	NFM_BOVIN
11	165.5	14.0	439	1	XP2_XENLA
12	162	13.8	5147	1	PCLO_HUMAN
13	160.5	13.6	992	1	EBN6_EBV
14	158	13.4	1274	1	ENAH_MOUSE
15	154	13.1	353	1	CCPA_ACEXY
16	154	13.1	4499	1	DYHA_CHIRE
17	152.5	12.9	5038	1	PCLO_MOUSE
18	152	12.9	721	1	YK82_MYCTU
19	151	12.8	88	1	H82_NEIGO
20	151	12.8	211	1	TUBB_SOLUT
21	150	12.7	352	1	ALGP_PSEAE
22	148.5	12.6	464	1	S3A2_HUMAN
23	147	12.5	1386	1	ZAP2_MOUSE
24	146.5	12.4	477	1	MAZ_MOUSE
25	145.5	12.4	361	1	IF35_MOUSE
26	145.5	12.4	1565	1	PAC_STRMU
27	145	12.3	475	1	S3A2_MOUSE
28	145	12.3	2167	1	SHK1_RAT
29	144.5	12.3	477	1	MAZ_HUMAN
30	142.5	12.1	497	1	WAS2_HUMAN
31	142.5	12.1	1083	1	T2D3_HUMAN
32	142	12.1	2161	1	SHK1_HUMAN
33	141.5	12.0	316	1	CDNC_HUMAN

34	141.5	12.0	465	1	FXD1_HUMAN	Q16676 homo sapien
35	140.5	11.9	852	1	WS14_HUMAN	Q9HP71 homo sapien
36	140	11.9	3421	1	TBGT_HSVB	P28955 equine hept
37	140	11.9	5085	1	PCLO_RAT	Q91966 ratuue novy
38	139.5	11.8	397	1	SEPL_MOUSE	Q62170 mus musculu
39	139.5	11.8	1157	1	BBC1_YEAST	P47068 saccharomyc
40	138.5	11.8	206	1	PRP1_MEDTR	Q43564 medicago tr
41	138.5	11.8	1183	1	DRPL_RAT	P54258 ratuue novy
42	138	11.7	376	1	FXL2_HUMAN	P58012 homo sapien
43	138	11.7	2517	1	NCR2_HUMAN	Q9Y618 h nuclear r
44	137	11.6	236	1	PRP_MEDSA	Q40358 medicago sa
45	136.5	11.6	1003	1	MBD6_HUMAN	Q96496 homo sapien

ALIGNMENTS

RESULT 1
GP1_CHLRE STANDARD; PRT; 555 AA.

AC Q9FQ6; Q03927; (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Vegetative cell wall protein gpl precursor (Hydroxyproline-rich glycoprotein 1).
GN GP1.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21159092; PubMed=11258910;
RA Ferris P.J., Woessner J.P., Welfenschmidt S., Kitz S., Drees J., Goodenough U.W.;
RT "Glycosylated polyproline II rods-with-kinks as a structural motif in plant hydroxyproline-rich glycoproteins";
RL Biochemistry 40:2978-2987(2001).
[2]
RN PARTIAL PRELIMINARY SEQUENCE FROM N.A.
RX MEDLINE=91017504; PubMed=1699225;
RA Adair W.S., Apt K.E.;
RT "Cell wall regeneration in Chlamydomonas: accumulation of mRNAs encoding cell wall hydroxyproline-rich glycoproteins";
RL Proc. Natl. Acad. Sci. U.S.A. 87:7355-7359(1990).
CC -!- FUNCTION: Major component of the outer cell wall w6 (crystalline) layer.
CC -!- SUBUNIT: Associates with GP2 and GP3.
CC -!- PTM: N-glycosylated and O-glycosylated.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC EMBL, AF309494; AAC5420.1; -.
CC EMBL, M58496; AAB69706.1; ALT_SEQ.
CC Glycosylated; Q9FQ6; -.
CC Interpro: IPR002965; P-rich extensin.
CC Interpro: IPR003882; PstII_extensin.
CC PRINTS: PRO1217; PRICHEXTENSIN.
CC PRINTS: PRO1218; PSTLTEXTENSIN.
CC Glycoprotein; Repeat; signal.
CC SIGNAL 1 29
FT CHAIN 30 555
FT DOMAIN 40 339
FT DOMAIN 259 379
FT CARBOHYD 399 399
FT CARBOHYD 455 455
FT POTENTIAL 49 X 5 AA APPROXIMATE PSEPK REPEATS.
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CABOARD 493      493      N-LINKED (GLCNAC... ) (POTENTIAL) .
SO SEQUENCE 555 AA; 54219 MM; 6A584A90465502F5 CXC64;

Query Match          15.3%; Score 180; DB 1; Length 555;
Best Local Similarity 25.7%; Pred. No. 0.00089;
Matches    56; Conservative   17; Mismatches    83; Indels    62; Gaps     5;

Cc      33 PLHPSLNTPGVIRNLNP-PLYYRPNVTNYSYRGNTTYDTDGHPSYPMWILTSQGPVVYHIR 91
Db      158 FVPSPSSPPVPPSPAPSPPTSPSSPPVPPSPAPPSPAPVPVPSB---APPSP----- 207
Cc      92 GFPIATQLNVPELPKGFPEPVPSRFSGNAAPAAPPAIAEPAAAPLTATPVAAEPAAG 151
Db      208 -----APPVPPSPAAPPSPS-----PAPSPPSPAAPPSPPAPSPFPVPPSPAP 252
Cc      152 APVAAEPAAE-----PVGAEPAAEAIPA 175
Db      253 SPAPPSPKPAPAPPPPSPPPPPPPPPANTMPMPPSBSPSPBPAPPTPTTSPSPSPS 312
Cc      176 AEPAAEAPVGVEPAAEESPAPAEPAATAKAADPHSPS 213
Db      313 PVPSAPAVPPSPAPAPSPAPSPBPSPAPPTSPSPSPS 350

RESULT 2
IF2_STIUA
ID_IF2 STIUA STANDARD; PRT; 1054 AA.
AC P55875;
DC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Translation initiation factor IF-2.
GN INPB.
OS Stigmatella aurantiaca.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
CC Cyctobacteriaceae; Cyctobacteraceae; Stigmatella.
OX NCBI_TaxID=41;
RN [1]
RP SEQUENCE FROM N.A.
RS STRAIN=Dm4;
RX MEDLINE=97234648; PubMed=9079922;
RA Bremard L., Laalami S., Deriard B., Cenatiempo Y.;
RT "Translation initiation factor Ifz of the myxobacterium Stigmatella
RL aurantiaca: presence of a single species with an unusual N-terminal
sequence." ;
RL J. Bacteriol. 179:2348-2355(1997).
CC CC -1- FUNCTION: One of the essential components for the initiation of
CC protein synthesis. Protects formylmethionyl-tRNA from spontaneous
CC hydrolysis and promotes its binding to the 30S ribosomal subunits.
CC Also involved in the hydrolysis of GTP during the formation of the
CC 70S ribosomal complex (By similarity).
CC CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC CC -1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on how
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
Cc      EMBL; X87940; CA61162.1; -.
DR      PIR; T43226; T43226.
DR      HSSP; P13551; IELO.
DR      HAMAP; MF_00100; -: 1.
DR      InterPro; IPR000795; EF_GTPbind.
DR      InterPro; IPR004161; EFTU_D2.
DR      InterPro; IPR000178; IF2_N.
DR      InterPro; IPR006847; IF2_N.
DR      InterPro; IPR005225; Small_GTP_Pfam; PF00009; GTP_EFTU; 1.
DR      Pfam; PF03144; GTP_EFTU_D2; 2.
```

```

DR PFam: PF04760; IF2_N; 2.
DR ProDom: PD186100; IF2; 1.
DR TIGRFAMs: TIGR00487; IF-2; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
DR PROSITE: PS01176; IF2; 1.
KW Initiation factor; Protein biosynthesis; GTP-binding.
FT DOMAIN 555 703 G-DOMAIN.
FT NP_BIND 561 568 GTP (BY SIMILARITY).
FT NP_BIND 607 611 GTP (BY SIMILARITY).
FT NP_BIND 661 664 GTP (BY SIMILARITY).
SQ SEQUENCE 1054 AA; 111323 MW; FFD519530B5D0669 CRC64;

Query Match 14.9%; Score 175; DB 1; Length 1054;
Best Local Similarity 38.3%; Pred. No. 0.0029;
Matches 49; Conservative 8; Mismatches 51; Indels 20; Gaps 4

OY 96 ATQLNVPPLPRGPFPPSPRFSAAPAAPPIAA--EPAAAPLVTAT--PVAAPRAG 151
DB 98 ASDVSSPSPSE-----VHEAGCAEAAASERVEEAAAVEPVAEAPRAAASEPAEAPKAT 152
OY 152 APVAEPAEAPVGAEPAAEA.PVAEAPAAEAPVGVTEPAEESP-----AEPAT 200
DB 153 APVAPEPVEAPKAAAPVAPEPTVEAPTEAPAAAPIAEAPTPAPTEVPTSGRRAS 212
OY 201 AKPAAPEP 208
DB 213 CRGAAPLP 220

RESULT 3
TEGU EBV STANDARD; PRT; 3149 AA.
ID TEGU EBV STANDARD; PRT; 3149 AA.
AC P03186;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Large tegument protein.
CN BPLF1
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
CX NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tuffnell P.S., Barrett B.G.;
RL "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211(1984).
CC -1- FUNCTION: TEGUMENT PROTEIN.
CC -1- FUNCTION: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
CC EHV-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboratio
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on it
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for Commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@sib.ch).
CC -----
CC EMBL: V01555; CAA24839.1; -.
DR PIR: G93065; Q0BE8.
DR InterPro: IPR006928; Herpes teg_N.
DR Pfam: PF04843; Herpes teg_N.1.
SQ SEQUENCE 3149 AA; 337954 MW; 3DD0C57658731D08 CRC64;

Query Match 14.9%; Score 175; DB 1; Length 3149;
Best Local Similarity 34.0%; Pred. No. 0.0071;
Matches 48; Conservative 19; Mismatches 40; Indels 34; Gaps 6

OY 105 PPRGPPF---VP-----PSRFSAAPAAAPAA-----PPIAAEPAAAPLVTITPVA 145

```

```

Db      290 PPVWLPEFRRIITPVYLRPLPKRSFTSDSEFPARRYSPAKTNSDP--SSPASADAPASAPAS 347
Qy      146 AEPAGAFAVAEPAEAFAVGAEPAAEAFAVAEPAFAVGVGAEPAAEBSPEAPATAKP-A 204
Db      348 AAPASAAAPASAAAPASAAAPASAAAPASAPSPPLTPIPIGLGHTGTGVAFTSPAPASGA 407
Qy      205 APE-----PHSPS 213
Db      408 APQTPKRRKGLGKDSFHKKPT 428

RESULT 4
TPM4_DROME STANDARD; PRT; 518 AA.
AC      P49455; P49456; Q24425; Q24426;
DT      01-FEB-1996 (Rel. 33, Created)
DT      01-FEB-1996 (Rel. 33, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Trpomyosin I, isoforms 33/34 (Trpomyosin II).
GN      TM1 OR TM11.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_taxid=7227;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORMS 33 AND 34).
RC      TISSUE=Embryo, and Pupae;
RX      MEDLINE=89127197; Pubmed=2851721;
RA      Hanke P.D., Scotti R.V.;
RT      "The Drosophila melanogaster tropomyosin II gene produces multiple
RT      proteins by use of alternative tissue-specific promoters and
RT      alternative splicing."
RL      Mol. Cell. Biol. 8:3591-3602(1988).
RN      [2]
RP      SEQUENCE FROM N.A. (ISOFORMS 33 AND 34).
RC      STRAIN=Oregon-R; TISSUE=Pupae;
RX      MEDLINE=87064486; Pubmed=3097506;
RA      Karlik C.C., Fyrberg E.A.;
RT      "Two Drosophila melanogaster tropomyosin genes: structural and
RT      functional aspects."
RL      Mol. Cell. Biol. 6:1965-1973(1986).
RN      [3]
RP      -1- ALTERNATIVE PRODUCTS:
RC      Event=Alternative splicing; Named isoforms=5;
RX      Comment=Additional isoforms seem to exist;
RA      Name=33; Synonyms=9c;
RA      IsoId=P49455-1; Sequence=Displayed;
RA      Name=Muscle; Synonyms=9D;
RA      IsoId=P06754-1; Sequence=External;
RA      Name=Non-muscle; Synonyms=Cytoskeletal;
RA      IsoId=P06754-2; Sequence=External;
RA      Name=9A;
RA      IsoId=P06754-3; Sequence=External;
RA      Name=34; Synonyms=9B;
RA      IsoId=P49455-2; Sequence=VSP 006623, VSP 006624, VSP 006625;
CC      -1- TISSUE SPECIFICITY: Both isoforms are only expressed in indirect
CC      flight muscles
CC      -1- DEVELOPMENTAL STAGE: Both isoforms are expressed during pupal and
CC      adult stages.
CC      -1- DOMAIN: THE MOLECULE IS IN A COILED COIL STRUCTURE. THE SEQUENCE
CC      EXHIBITS A PROMINENT SEVEN-RESIDUES PERIODICITY.
CC      -1- SIMILARITY: BELONGS TO THE TRPOMYOSIN FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X76208; CAA53800.1; -.

```

Query Match	Best Local Similarity	Matches	Score 174;	DB 1;	Length 518;	Pred. No. 0.0018;	Mismatches 107;	Indels 58;	Gaps 7;
17	KRRPPPTGENDNDGCHLHPSELINPYGRNLPPLUYAPVNT	58							
264	KERICMIGDSLDEAFVLLIKGLEFPNWRKRPPLKPLPTTPEELAMEERAAAEAA	323							
59	-----VPSYGNITYTDTGLPSYMWLT--SPGFPVYHI---R	91							
324	AAAEAAEAAMAAAGADGAPAAAPGEEKAPAKEPTPEKPTPPPPPPPPREYSIDLPE	383							
92	GFPLATQUNVPLPPRGFPFVPPGFRFSAAAP--AAPPT--AAEPALAAPLTATVAAEP	148							
384	GAEPVYKYNPEPPPGSEPEVPFAEGEAAAPAAAGAAPPAAGAAPPAAGAAPPAAGAAP	443							
149	AAAGAPVAAEPAAEAPVGAEPAAEAPV-----AAEPAAEAPVGVVPAAS-----EPSPAP	198							
444	AEGAAPAAEGAAPPAADGAPAAEPAAEAAPADAAAAPAAEAAPAAEPAAEAAPAAEPAAEP	503							
199	ATAKPAAPBPHSPS	213							

DB 504 ABAAPAAEGCAPA 518

RESULT 5
CA36 HUMAN STANDARD; PRT; 3176 AA.

AC P12111; Q16501; 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Collagen alpha 3 (VI) chain precursor.
GN COL6A3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Fibroblast;
RX MEDLINE=90151612; PubMed=1689238;
RA Chu M.-L., Zhang R.-Z., Pan T.-C., Stokes D., Conway D., Kuo H.-J.,
RA Glaville R., Mayer U., Mann K., Deutzmann R., Timpl R.;
RT "Mosaic structure of globular domains in the human type VI collagen
alpha 3 chain: similarity to von Willebrand factor, fibronectin,"
RT actin, salivary proteins and apocytin type protease inhibitors,"
RL EMBO J. 9:385-393(1990).
RN [2]
RP REVISIONS.
RA Chu M.-L.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 2038-2373 FROM N.A.
RX MEDLINE=89066644; PubMed=3198591;
RA Chu M.-L., Conway D., Pan T.-C., Baldwin C., Mann K., Deutzmann R.,
RA Timpl R.;
RT "Amino acid sequence of the triple-helical domain of human collagen
type VI,"
RL J. Biol. Chem. 263:18601-18606(1988).
RN [4]
RP SEQUENCE OF 2092-2157 FROM N.A.
RX MEDLINE=88029444; PubMed=3665927;
RA Chu M.-L., Mann K., Deutzmann R., Pribula-Conway D.,
RA Han-Chen C.-C., Bernard M.P., Timpl R.;
RT "Characterization of three constituent chains of collagen type VI by
peptide sequences and cDNA clones,"
RL Eur. J. Biochem. 168:309-317(1987).
RN [5]
RP SEQUENCE OF 2092-2151 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=88161046; PubMed=3348212;
RA Well D., Mettel M.-G., Passage E., van Cong N., Pribula-Conway D.,
RA Mann K., Deutzmann R., Timpl R., Chu M.-L.;
RT "Cloning and chromosomal localization of human gene encoding the
three chains of type VI collagen,"
RL Am. J. Hum. Genet. 42:435-445(1988).
RN [6]
RP SEQUENCE OF 32-236 FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=93054780; PubMed=1339440;
RA Zanussi S., Doliana R., Segat D., Bonaldo P., Colombatti A.;
RT "The human type VI collagen gene. mRNA and protein variants of the
alpha 3 chain generated by alternative splicing of an additional 5-end
exon,"
RL J. Biol. Chem. 267:24082-24089(1992).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 3107-3164.
RX MEDLINE=95182468; PubMed=7533217;
RA Arnoux B., Merigues K., Saludjian P., Norris F., Norris K., Bjoern S.,
RA Olsen O., Petersen L., Ducruix A.;
RT "The 1.6 A structure of Kunitz-type domain from the alpha 3 chain of
human type VI collagen,"
RL J. Mol. Biol. 246:609-617(1995).
RN [8]

RP STRUCTURE BY NMR OF 3102-3164.
RX MEDLINE=96398604; PubMed=8805527;
RA Zweckstetter M., Czisch M., Mayer U., Chu M.-L., Zinth W., Timpl R.,
RA Holak T.A.;
RT "Structure and multiple conformations of the Kunitz-type domain from
human type VI collagen alpha3 (VI) chain in solution,"
RL Structure 4:195-209(1996).
RN [9]
RP STRUCTURE BY NMR OF 3107-3164.
RX MEDLINE=97410331; PubMed=9265624;
RA Soerensen M.D., Bjoern S., Norris K., Olsen O., Petersen L.,
RA James T.L., Led J.J.;
RT "Solution structure and backbone dynamics of the human alpha3-chain
type VI collagen C-terminal Kunitz domain,"
RL Biochemistry 36:10439-10450(1997).
RN [10]
RP DISEASE.
RX MEDLINE=21987636; PubMed=11992252;
RA Demir E., Sabatelli P., Allamand V., Ferreiro A., Moghadaszadeh B.,
RA Makrelouf M., Topaloglu H., Echenne B., Merlini L., Guicheney P.;
RT "Mutations in COL6A3 cause severe and mild phenotypes of Ullrich
congenital muscular dystrophy,"
RL Am. J. Hum. Genet. 70:1446-1458(2002).
RN [11]
RP VARIANT BM GLU-1679, AND VARIANT HIS-2831.
RX MEDLINE=98204804; PubMed=9536084;
RA Pan T.-C., Zhang R.-Z., Pericak-Vance M.A., Tandan R., Fries T.,
RA Stajich J.M., Viles K., Vance J.M., Chu M.-L., Speer M.C.;
RT "Missense mutation in a von Willebrand factor type A domain of the
alpha 3 (VI) collagen gene (COL6A3) in a family with Bethlem
myopathy,"
RL Hum. Mol. Genet. 7:807-812(1998).
RN [12]
RP FUNCTION: COLLAGEN VI ACTS AS A CELL-BINDING PROTEIN.
CC -1- SUBUNIT: TRIMERS COMPOSED OF THREE DIFFERENT CHAINS: ALPHA 1 (VI),
CC ALPHA 2 (VI), AND ALPHA 3 (VI).
CC -1- ALTERNATIVE PRODUCTS:
CC Name=1;
CC IsoId=P12111-1, Sequence=Displayed;
CC Name=2;
CC IsoId=P12111-2, Sequence=VSP 001172;
CC -1- PTM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -1- DISEASE: Defects in COL6A3 are a cause of Bethlem myopathy (BM)
CC [MIM:158810]. BM is a rare autosomal dominant proximal myopathy
CC characterized by early childhood onset (complete penetrance by the
CC age of 5) and joint contractures most frequently affecting the
CC elbows and ankles.
CC -1- DISEASE: Defects in COL6A3 are a cause of Ullrich congenital
CC muscular dystrophy (UCMD) [MIM:254090]; also known as Ullrich
CC scleroatonic muscular dystrophy, an autosomal recessive congenital
CC myopathy. UCMD is characterized by muscle weakness and multiple
CC joint contractures, generally noted at birth or early infancy. The
CC clinical course is more severe than in Bethlem myopathy.
CC -1- SIMILARITY: Contains 1 BPT/Kunitz inhibitor domain.
CC -1- SIMILARITY: Contains 1 fibronectin type III domain.
CC -1- SIMILARITY: Contains 12 WFPA domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52022; CAA36267.1; -;
DR EMBL; X06186; CAA29557.1; -;
DR EMBL; M20778; -; NOT_ANNOTATED_CDS.
DR EMBL; M27449; AAB52057.1; -;
DR EMBL; S49432; AAB24261.1; -;
DR PIR; AS9140; CGHJ3A.
DR PDB; 1KNT; 01-NOV-94.

```

DR PDB: 2KNT; 15-MAY-97.
DR PDB: 1KUN; 12-NOV-97.
DR PDB: 1KTH; 28-AUG-02.
DR Genew; HGNC:2213; COL6A3.
DR MIM; 120250; -.
DR MIM; 158810; -.
DR MIM; 254090; -.
DR GO; GO:0005589; C:collagen type VI; TAS.
DR GO; GO:0005202; F:collagen; TAS.
DR GO; GO:0007517; P:muscle development; TAS.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR002223; Kunitz_Bpt1.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01391; Collagen_5.
DR Pfam; PF00014; Kunitz_Bpt1; 1.
DR Pfam; PF00092; vwa; 11.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000007; Clg helix; 2.
DR ProDom; PD000222; Kunitz_Bpt1; 1.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00131; KU; 1.
DR SMART; SM00327; VWA; 12.
DR PROSITE; PS00280; BPT1_KUNITZ_1; 1.
DR PROSITE; PS50279; BPT1_KUNITZ_2; 1.
DR PROSITE; PS50234; VWFA; 12.
DR ExTracellul; Connective tissue; Repeat; Hydroxylation;
DR Glycoprotein; Cell adhesion; Serine protease inhibitor; Collagen;
DR Signal; 3D-structure; Disease mutation; Polymorphism;
DR Alternative splicing.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 3176 COLLAGEN ALPHA 3(VI) CHAIN.
FT DOMAIN 26 2038 NONHELICAL REGION.
FT DOMAIN 2039 2375 TRIPLE-HELICAL REGION.
FT DOMAIN 2376 3176 NONHELICAL REGION.
FT DOMAIN 39 213 VWFA 1.
FT DOMAIN 242 419 VWFA 2.
FT DOMAIN 445 620 VWFA 3.
FT DOMAIN 639 816 VWFA 4.
FT DOMAIN 837 1009 VWFA 5.
FT DOMAIN 1029 1205 VWFA 6.
FT DOMAIN 1233 1404 VWFA 7.
FT DOMAIN 1436 1609 VWFA 8.
FT DOMAIN 1639 1812 VWFA 9.
FT DOMAIN 1838 2024 VWFA 10.
FT DOMAIN 2042 2581 VWFA 11.
FT DOMAIN 2619 2815 VWFA 12.
FT DOMAIN 2987 3076 FIBRONECTIN TYPE-III.
FT DOMAIN 3107 3176 BPT1/KUNITZ INHIBITOR.
FT SITE 2040 2042 CELL ATTACHMENT SITE.
FT SITE 2136 2138 CELL ATTACHMENT SITE.
FT SITE 2148 2150 CELL ATTACHMENT SITE.
FT SITE 2154 2156 CELL ATTACHMENT SITE.
FT SITE 2370 2372 CELL ATTACHMENT SITE.
FT ACT SITE 3121 3122 REACTIVE BOND.
FT DISULFID 3111 3161
FT DISULFID 3110 3144
FT DISULFID 3120 3144
FT CAROHYD 202 202
FT CAROHYD 251 251
FT CAROHYD 251 251

```

```

Query Match 14.4%; Score 170; DB 1; Length 3176;
Best Local Similarity 31.8%; Pred. No. 0.014;
Matches 57; Conservative 16; Mismatches 56; Indels 50; Gaps 9;

```

```

Qy 35 HPSLIPGIRLPL---PPLYRPNVTGSGNTYTDGLPSYMWILSPGFPVYHNR 91
Db 2853 HKQVAVPNNVSSPSNPTTKKPVYTT-----KPVYTTKP-----VYTTTKP----- 2896
Qy 92 GEPPLATOLNVPPLPRGPFVPPSPRFFSAAPAA-APPIAEPAAAPLATATPVAEPAA 150
Db 2897 -----VTIINQPSVKF-----AAAKPAPAKPVAAPVATKATVAP----- 2932

```

```

Qy 151 GAPVAEPAAEA-PVGAEPAAEPVAAEPAAEPVAAEPAAEPAAEPAAEPAAEP 208
Db 2933 --PVAVKATAPAKPVAAPKPAAPVAPRA---AAAKPAPAKPVAAPVATKATVAP 2986

```

RESULT 6

ID ENAH_MOUSE STANDARD. PRT. 802 AA.

AC Q03173; P70430; P70431; P70432; P70433;

DT 01-OCT-1996 (Rel. 34, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Enabled protein homolog (NPC derived proline-rich protein 1) (NDPP-1).

GN ENAH OR MENA OR NDP1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Brain;

RX MEDLINE=93041923; PubMed=1420303;

RA Sazuka T., Tomooka Y., Kathju S., Ikawa Y., Noda M., Kumar S.;

RT "Identification of a developmentally regulated gene in the mouse central nervous system which encodes a novel proline rich protein.";

RL Biochim. Biophys. Acta 1132:240-248(1992).

[2]

RP SEQUENCE FROM N.A. (ISOFORMS 2; 3; 4 AND 5), FUNCTION, AND SUBCELLULAR LOCATION.

RC TISSUE=Brain;

RX MEDLINE=97015079; PubMed=8861907;

RA Gertler F.B., Niebur K., Reinhard M., Wehland J., Soriano P.;

RT "Mena, a relative of VASP and Drosophila Enabled, is implicated in the control of microfilament dynamics.";

RL Cell 87:227-239(1996).

[3]

RP FUNCTION, AND SUBUNIT.

RX MEDLINE=9916667; PubMed=10069337;

RA Lanier L.M., Gates M.A., Witke W., Menzies A.S., Wehman A.M., Macklis J.D., Kwiatkowski D., Soriano P., Gertler F.B.;

RT "Mena is required for neurulation and commissure formation.";

RL Neuron 22:313-325(1999).

CC -1- FUNCTION: May be involved in microfilament assembly and cell motility. Induces the formation of F-actin rich outgrowths in fibroblasts. Required for neurulation and commissure formation.

CC -1- SUBUNIT: Binds profilin.

CC -1- SUBCELLULAR LOCATION: Localized to focal adhesions and, to a lesser extent, leading edges and stress fibers.

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=5;

CC Name=5; Synonyms=Mena+++;

CC IsoId=Q03173-1; Sequence=Displayed;

CC Name=1;

CC IsoId=Q03173-2; Sequence=VSP_007255, VSP_007256;

CC Name=2; Synonyms=Mena;

CC IsoId=Q03173-3; Sequence=VSP_007259, VSP_007260;

CC Name=3; Synonyms=Mena+;

CC IsoId=Q03173-4; Sequence=VSP_007259;

CC Name=4; Synonyms=Mena++;

CC IsoId=Q03173-5; Sequence=VSP_007257, VSP_007258;

CC -1- TISSUE SPECIFICITY: IN THE HEART AND TESTIS AND LESS SO IN THE LUNG, SKELETAL MUSCLE, KIDNEY, PANCREAS AND BRAIN.

CC -1- SIMILARITY: Contains 1 WH1 domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; D10727; BA01570.1; -.


```

DR  FLYBase: FBgn0015288; RPL22.
DR  InterPro: IPR002671; Ribosomal_L22e.
DR  Pfam: PF01776; Ribosomal_L22e; 1.
DR  ProDom: PD007306; Ribosomal_L22e; 1.
KW  Ribosomal protein.
FT  DOMAIN 24 31 POLY-ALA.
FT  DOMAIN 46 50 POLY-ALA.
FT  DOMAIN 65 70 POLY-ALA.
FT  DOMAIN 93 98 POLY-ALA.
FT  DOMAIN 103 112 POLY-ALA.
FT  DOMAIN 136 152 POLY-ALA.
FT  DOMAIN 185 188 POLY-LYS.
FT  DOMAIN 292 299 ASP/GLU-RICH (HIGHLY ACIDIC).
SO  SEQUENCE 239 AA; 30610 MW; 46A99005610E4EB0 CRC64;

Query Match 14.3%; Score 169; DB 1; Length 239;
Best Local Similarity 44.1%; Pred. No. 0.0022;
Matches 49; Conservative 6; Mismatches 34; Indels 22; Gaps 5

Cc 121 AAAPAPPIAAEPAAADLVLTAT-----PVAAEPPAGAPVAEPPAAEPVGAEPAAEP 173
Cc ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cc 65 AAAAAAKPPAAKPPAAKSKDAGKAPAAAPAAKPKAKAAAAAPAKPAKKAASTP 124
Cc ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

Cy 174 VAAEPA-----AEAPGVPEPAAEPSAPAPATKPPAPEPH-----PSPS 213
Cy ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cy 125 AAAPPAKKAAPAKAAAPAAAPAA--PAAAPAAVAPK-APKPKAKAAPAPS 171
Cy ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
CPN_DROME CPN_DROME STANDARD: PRT: 865 AA.
AC Q02910;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Calphoctin.
DE CPN OR CAP.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Empidoidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=93165729; PubMed=8094559;
RA Martin J.H., Benzer S., Rudnicka M., Miller C.A.;
RT "Calphoctin: a Drosophila photoreceptor cell calcium-binding protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:1531-1535(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=93165730; PubMed=8434015;
RA Ballinger D.G., Xue N., Harshman K.D.;
RT "A Drosophila photoreceptor cell-specific protein, calphoctin, binds
RT calcium and contains a leucine zipper.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:1536-1540(1993).
CC -1- FUNCTION: Might function as a calcium-sequestering "sponge" to
CC regulate the amount of free cytoplasmic calcium. It binds 0.3 mole
CC of Ca(2+) per mole of protein.
CC -1- SUBUNIT: Homodimer (Probable).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; HYPODENSE COMPARTMENT.
CC -1- TISSUE SPECIFICITY: SOMA AND AXONS OF PHOTORECEPTOR CELLS OF
CC COMPOUND EYES AND OCELLI.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN PHOTORECEPTOR CELL
CC DEVELOPMENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@lab-81b.ch).
CC -----
DR EMBL, L02111; AAA28405.1; -.
DR EMBL, L05080; AAA28420.1; -.
DR PIR; A47282; A47282.
DR PIR; A47283; A47283.
DR FlyBase; FBpp0010218; Cpn.
DR GO; 00005509; F:calcium ion binding activity; IDA.
DR KW Calcium-binding.
FT FT CONFLICT 36 36 A -> AAVAPAVA (IN REF. 2) .
FT FT CONFLICT 43 43 I -> T (IN REF. 2) .
FT FT CONFLICT 64 64 I -> V (IN REF. 2) .
FT FT CONFLICT 76 76 T -> A (IN REF. 2) .
FT FT CONFLICT 100 100 P -> PP (IN REF. 2) .
FT FT CONFLICT 126 127 VQ -> AP (IN REF. 2) .
FT FT CONFLICT 154 154 I -> V (IN REF. 2) .
FT FT CONFLICT 160 160 S -> T (IN REF. 2) .
FT FT CONFLICT 534 534 A -> E (IN REF. 2) .
FT FT CONFLICT 699 699 I -> T (IN REF. 2) .
FT FT CONFLICT 703 703 V -> L (IN REF. 2) .
FT FT CONFLICT 721 721 D -> E (IN REF. 2) .
SQ SEQUENCE 865 AA; 84781 MW; 2110417E0B087CFE CRC64;

Query Match 14.3%; Score 169; DB 1; Length 865;
Best Local Similarity 29.5%; Pred. No. 0.0053;
Matches 70; Conservative 14; Mismatches 75; Indels 78; Gaps 11

QY 55 PVNTVPSYGNTYDITDGLPSYPMILTSBGFYVYHINRGPPLAT--QLNVPELPGRGPEF 111
DB 42 PIATVPVAPPTTASV-----QPATVITPAPAPIAAASGTVPASVAPVVAAPTPTASV 97
QY 112 VPP--SRFESAAPAAPAPPIAEPBA-----AAAPLTAT- PVAAE-----PAA----- 150
DB 98 STPVAVAQIPVAVSAPVAPVAPVAPVATPTPVQVIPVAPVATATPPVVASAPPTPAVTPVISV 157
QY 151 -----GAPVAAEPB-----AAPVGAE-----PAAEAPVA 175
DB 158 IASPPVVPANTTVEVAPVAPVAPVAPVAPVAVLAPVAPVAPVAAETPAPPPVAETIPVA 217
QY 176 AEPAAEPV-----GEPAAEBSPSPATKAPPAEPHPSS--PSLEQA 217
DB 218 TTFPCVAPLIEVSVAATKPLAAAEFPVVAAPPAETTPVVAAPAAASPHYSVAPAVETA 274

RESULT 9
CY1_PARDE STANDARD; PRT; 450 AA.
ID CY1 PARDE
AC P15627;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome c1 precursor.
GN PUTC.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodobacteraceae; Paracoccus.
OX NCBI_TaxId=266;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88007612; PubMed=2820981;
RA Kuroweki B., Ludwig B.;
RT "The genes of the Paracoccus denitrificans bcl complex. Nucleotide
RT sequence and homologues between bacterial and mitochondrial
RT subunits."
RL J. Biol. Chem. 262:13805-13811 (1987).
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS. C1 FUNCTIONS AS AN ELECTRON DONOR TO
CC CYTOCHROME C.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIBESKE PROTEIN.
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; M17522; AAA2573.1; -
 CC EMBL; X05799; CAA29245.1; -
 CC PIR; C29413; C29413.
 CC InterPro; IPR0002326; Cyt C1.
 CC InterPro; IPR000345; CytC_heme_bind.
 CC Pfam; PF02167; Cytochrome_C1.
 CC PRINTS; PR00603; CYTOCHROME_C1.
 CC DR PROSITE; PS00190; CYTOCHROME_C; 1.
 CC K W Electron transport; Respiratory chain; Heme; Transmembrane; Signal.
 CC FT SIGNAL 1 21
 CC FT CHAIN 22 450 CYTOCHROME C1.
 CC FT BINDING 245 245 HEME (COVALENT).
 CC FT BINDING 248 248 HEME (COVALENT).
 CC FT METAL 249 249 IRON (HEME AXIAL LIGAND).
 CC FT METAL 373 373 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC FT TRANSMEM 421 435 ANCHORS TO THE MEMBRANE (POTENTIAL).
 CC SQ SEQUENCE 450 AA; 46874 MW; CLEA8BE48FE7603 CRC64;

Query Match 14.2%; Score 167.5; DB 1; Length 450;
 Best Local Similarity 42.6%; Pred. No. 0.0037;
 Matches 58; Conservative 8; Mismatches 35; Indels 35; Gaps 8;

QY 114 PERPFSMAAPAPPIAEPAPAA--APLATPVAA-----EPAGAPVA-----AEPAA 159
 DB 48 PAAADTAPAAEADEPAAEAEGAEVTEBPATTPAEPEEPAPATTEBPDAEAPAA 107

QY 160 -----AAEAVGAPAAEAPVAAEPAAEPAAEPVAAEPAAEP--ATAKPA 205
 DB 108 ABEAQATTEAPAEPEAPAEAPAEAP--AEPPAADAN--AEBAAPAEAPAEPEAPAAEPAA 164

QY 206 PEP---HPSPSLEQA 217
 DB 165 EEPATEEAPAEAPAA 180

RESULT 10
 NFM_BOVIN STANDARD; PRT; 810 AA.
 AC 077788;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet M protein (160 kDa neurofilament protein)
 GN NEF3 OR NEFM OR NFM.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NC NCB1_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hill W.D., Zhang L., Balin B.J., Sprinkle T.J., Spicer K.,
 RA Gearhart D.A.;
 RT "The bovine neurofilament M subunit has a novel set of KSP repeats
 RT normally restricted to NF-H.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- PTM: THERE ARE A NUMBER OF REPEATS OF THE TRIPEPTIDE K-S-P, NFM IS
 CC PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS
 CC THOUGHT THAT PHOSPHORYLATION OF NFM RESULTS IN THE FORMATION OF
 CC INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE
 CC OF AXONAL CALIBER (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING

CC OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE
 CC LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND
 CC COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; AF091342; AAC36357.1; -
 CC InterPro; IPR001664; IF.
 CC InterPro; IPR002957; Keratin_I.
 CC Pfam; PF00038; filament; 1.
 CC DR PRINTS; PR01248; TYPE1KERATIN.
 CC DR PROSITE; PS00226; IF; 1.
 CC KM Intermediate filament; Coiled coil; Neutrone; Phosphorylation.
 CC FT NON_TER 1
 CC FT DOMAIN 1 296 ROD.
 CC FT DOMAIN 297 810 TAIL.
 CC FT DOMAIN 297 810 COIL 1A.
 CC FT DOMAIN 21 33 LINKER 1.
 CC FT DOMAIN 34 132 COIL 1B.
 CC FT DOMAIN 133 149 LINKER 12.
 CC FT DOMAIN 150 171 COIL 2A.
 CC FT DOMAIN 172 175 LINKER 2.
 CC FT DOMAIN 176 296 COIL 2B.
 CC FT DOMAIN 503 582 8 X 13 AA TANDEM REPEATS.
 CC SQ SEQUENCE 810 AA; 90799 MW; B8477D85560AC3F6 CRC64;

Query Match 14.1%; Score 166; DB 1; Length 810;
 Best Local Similarity 35.1%; Pred. No. 0.0073;
 Matches 34; Conservative 12; Mismatches 51; Indels 0; Gaps 0;

QY 112 VPPSRPFSMAAPAPPIAEPAPAA--APLATPVAAEPAAEPAAEPVAAEPVGAEPAAE 171
 DB 484 VAPKEELAAEAVEREKEKASPVASPTTKSPTAKSPKASPEAKSPTKASPTAKSPVAK 543

QY 172 APVAAEPAAEAPVGAEPAAEPSPAPAPAPAPAP 208
 DB 544 SPTAKSPKASPEAKSPTKASPTAKSPKASPPAPSP 580

RESULT 11
 XP2_XENLA STANDARD; PRT; 439 AA.
 AC P17437; Q08944;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Skin secretory protein XP2 precursor (APBG protein).
 GN Xenopus laevis (African clawed frog).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE OF 1-25 AND 344-439 FROM N.A. (ISOFORM 2).
 RX MEDLINE=92332564; PubMed=1629230;
 RA Hauser F., Roeben C., Hofmann W.;
 RT "XP2, a new member of the P-domain peptide family of potential growth
 RT factors, is synthesized in Xenopus laevis skin.";
 RL J. Biol. Chem. 267:14451-14455(1992).
 RN [2]
 RP SEQUENCE OF 3-439 FROM N.A. (ISOFORM 1).
 RX TISSUE=Skin;
 RC MEDLINE=90127399; PubMed=2298293;
 RA Gmachl M., Berger H., Thalhammer J., Krell G.;
 RT "Dermal glands of Xenopus laevis contain a polypeptide with a highly

SQ SEQUENCE 353 AA; 37396 MW; 287456CAE5E7FB6F CRC64;

Query Match	Score	DB 1	Length	353
13.1%	154			

Best Local Similarity	29.9%	Pred. No. 0.017;	
Matches	52;	Conservative	9;
		Mismatches	73;
		Indels	40;
		Gaps	7.

Matches	52;	Conservative	9;	Mismatches	73;	Indels	40;	Gaps	7;
---------	-----	--------------	----	------------	-----	--------	-----	------	----

QY 53 YRPVNTVPSYPGNTYTTDTGLPSYPWILTSFGFPYVYHIRGFPLATQLNVPLPRGPFV 112

```

Db      37 YRP-----FVDRSFDVTGVE-----AVERHFDQAEH-----DTAVEBQVTPAP----- 75

```

QY 113 PPSRFFSAAAPAPPIAAEPAAAPLTATPVAAEPAGAPVAAEPAAEPVGAEPAAEA 172

```
Db      76 -----QIAVAPRPVDPRAIVTETAPRPV--VVSAPVTYEPRAAAVPAEPVQEA 127
```

QY 173 PVAAEPAAEPVGVTEPAAEESPAP-----ATAKPAAPEPHPSPSLEQA 217

Db 128 PVQAAPVPPAP--VPPIAEQAPPAAPDPASVPPYANVAAPVPPDPAPVTPAPQA 179

Search completed: January 15, 2004, 15:12:28
Job time : 19 secs

Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:11:09 ; Search time 21 Seconds

(without alignments)
1002.902 Million cell updates/sec

Title: US-09-923-236-2

Perfect score: 1178
Sequence: 1 MKLLIMACIVCVAFARAKRRF.....TAKPAAPRPHPSLEQANQ 219

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	206.5	17.5	607	2	S27776
2	199	16.9	391	2	S27850
3	193	16.4	3534	2	T42567
4	188.5	16.0	1108	2	T18353
5	188	16.0	839	2	F75518
6	186.5	15.8	395	2	H75457
7	183	15.5	744	2	B86255
8	179	15.2	207	2	A40215
9	175	14.9	1054	2	T43226
10	175	14.9	3149	1	Q08B8
11	174.5	14.8	1611	2	T18826
12	173.5	14.7	786	2	T01456
13	172.5	14.6	846	2	S52418
14	171.5	14.6	438	2	G87675
15	171	14.5	544	2	T17547
16	170.5	14.5	601	2	S27777
17	170	14.4	548	2	E70546
18	170	14.4	873	2	A47283
19	170	14.4	3176	1	CGH03A
20	169	14.3	865	2	A47282
21	169	14.3	1099	2	T18357
22	168.5	14.3	217	2	S29309
23	168.5	14.3	309	2	G83013
24	167.5	14.2	450	1	C29413
25	165.5	14.0	416	1	SKXLAG
26	165.5	14.0	416	2	T34279
27	165.5	14.0	801	2	T29018
28	165.5	14.0	1015	2	JC6552
29	165	14.0	214	2	T10737

30	163.5	13.9	214	2	T09854	proline-rich cell
31	162	13.8	98	2	H81072	hypothetical prote
32	162	13.8	428	2	T24769	hypothetical prote
33	162	13.8	460	2	T33110	hypothetical prote
34	160.5	13.6	352	2	A36128	regulatory protein
35	160.5	13.6	872	2	S33015	hypothetical prote
36	160.5	13.6	992	2	A31666	hypothetical prote
37	158	13.4	340	2	A35610	regulatory protein
38	158	13.4	1274	2	T37193	enamelin matrix pr
39	157	13.3	678	2	A75580	hypothetical prote
40	156	13.2	581	2	T22341	hypothetical prote
41	156	13.2	4957	2	T03455	ALR protein - huma
42	155.5	13.2	93	2	S04157	outer membrane pro
43	155	13.2	574	2	T43556	Wiskott-Aldrich sy
44	155	13.2	574	2	T38819	wiskott-aldrich sy
45	154.5	13.1	878	2	T21621	hypothetical prote

ALIGNMENTS

RESULT 1

S27776
80K protein (allele C1B) - Babesia bovis

C.Species: Babesia bovis

C.Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Sep-1997

C.Accession: S27776

R.Datymple, B.P.; Peters, J.M.

Submitted to the EMBL Data Library, May 1992

A.Description: Sequence of cDNA clones of a babesia bovis gene isolated using sera from

A.Reference number: S27776

A.Accession: S27776

A.Molecule type: mRNA

A.Residues: 1-607 <DAL>

A.Cross-references: EMBL:M93126; NID:g155862; PID:g155863

Query Match

Best Local Similarity 17.5% Score 206.5; DB 2; Length 607;

Matches 46; Conservative 7; Mismatches 35; Indels 11; Gaps 2;

QY 128 PIAAEPAAPADLTATPVAEPAPACAPVAEPAAEPVGAEPAAEPVGAEPVGE 187
DB 329 PIVEEPVAEPPIVEEPPIVEEPPIVEEPVAEPVAEPVAEPVAEPVAEPPIVEE 388

QY 188 PAEEP-----SPAEPATAKPA--APEBPSPSIE 215
DB 389 PVAEPVAEPBPVCEPAPETPAKPAETPAKPAETPAKPAE 427

RESULT 2

S27850
surface antigen Tc13 - Trypanosoma cruzi (fragment)

C.Species: Trypanosoma cruzi

C.Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 11-Jan-2000

C.Accession: S27850

R.Campetella, O.E.

Submitted to the EMBL Data Library, April 1992

A.Reference number: S27850

A.Accession: S27850

A.Molecule type: DNA

A.Residues: 1-391 <CAM>

A.Cross-references: EMBL:M92046; NID:g161955; PID:g161956

C.Superfamily: histone H1

Query Match

Best Local Similarity 16.9% Score 199; DB 2; Length 391;

Matches 42; Conservative 14; Mismatches 40; Indels 0; Gaps 0;

QY 113 PPSRFSAAPAAPPIAEPAAAPLTATVAEPAAAGVAAEPAAEPVGAEPAAEA 172
DB 5 PKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAE 64

QY 173 PVAEPAAEPVGEVPAEPAAEPSPAPPAEPATAKPAAPBP 208

A:Molecule type: DNA
A:Residues: 1-395 <WHI>
A:Cross-references: GB:AB001946; GB:AB000513; NID:g6458655; PIDN:AAFI0516.1; PID:g645866
A:Experimental source: strain RI
C:Genetics:
A:Gene: DR0938
A:Map position: 1

Query Match 15.8%; Score 186.5; DB 2; Length 395;
Best Local Similarity 35.0%; Pred. No. 0.0001;
Matches 57; Conservative 11; Mismatches 64; Indels 31; Gaps 9;

QY 60 PSYF-GNTYDTGLSYFPIILTSFGSPYVYHGRFPLATQLWVPLPPGPFVPSRPF 118
DB 53 PAEPVGAQPTTSTIVSGAVVADSPQT---AAPAQQTAKGKIPPA--PAAP----- 102
QY 119 SAAAAPAAP-----IAEPMAAPLTATPVAEPAAAGAPVAEPAAEPVGAEPAA 170
DB 103 ----APKIPPLPPLPRRMPAPPTPRPVEETTTAASPTQPPVQAAPATQPTPTQTAA 158
QY 171 EAPVAAEPAAEPVGEPA--EEGSPA--EPAT-AKPAPEP 208
DB 159 QAPATQPPATPAP---EPAAAPAPAPATTPPATPTPEPATTEPAAPAP 198

RESULT 7

hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
A:Accession: E66255
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Nasser, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
Ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E66255
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-744 <STO>
A:Cross-references: GB:AB005172; NID:g3157926; PIDN:AACT1609.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 15.5%; Score 183; DB 2; Length 744;
Best Local Similarity 28.9%; Pred. No. 0.00031;
Matches 55; Conservative 14; Mismatches 93; Indels 28; Gaps 6;

QY 33 PLHPLNIPYGRINLPPLIYRPVNTVPSYRGNTYTDGLSPY-----PILTSRGP 85
DB 461 PPSPPPPPPYVSSPPPPYV---SSPPPPYVSSPPPPYVSSPPPPYVSSPPPP 516
QY 86 YUHHRGFPLATQLWVPLPRGPFVPSRFFSAAAAPAA--PPIAEPAAALATATP 143
DB 517 YYY-----SSPPPP---PSPPPCPSSPPPPVYVAVVQSPPPSPSVYYYP 562
QY 144 VAAEPAAAGAPVAEPAAEPVGAEPAAEPVAAEPAAEPVGEPAEPSPSAEPATAP 203
DB 563 VTQSPPPSPSVYVPPVNTSPPPSPSVYVPPVNTSPPPSPSVYVPPVNTSPPPSPSVYVPP 622
QY 204 AAPRPPSPS 213
DB 623 VTPSP-PPPS 631

RESULT 8

AA0215

TCD antigen - Trypanosoma cruzi

C:Species: Trypanosoma cruzi
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
A:Accession: AA0215
R:Burns Jr., J.M.; Shreffler, W.G.; Roeman, D.E.; Sleath, P.R.; March, C.J.; Reed, S.G.
Proc. Natl. Acad. Sci. U.S.A. 89, 1239-1243, 1992
A:Title: Identification and synthesis of a major conserved antigenic epitope of Trypanos
A:Reference number: AA0215; MUID:92159024; PMID:1371355
A:Accession: AA0215
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-207 <BUR>
A>Note: sequence extracted from NCBI backbone (NCBIN:82113, NCBI:P:82117)
C:Superfamily: histone H1

Query Match 15.2%; Score 179; DB 2; Length 207;
Best Local Similarity 41.0%; Pred. No. 0.00016;
Matches 43; Conservative 11; Mismatches 45; Indels 6; Gaps 2;

QY 114 PSRFFSAAAAPAPPIAEPAAAPLTATPVAEPAAAGAPVAEPAAEPVGAEPAAEP 173
DB 5 PAEPKSAEPKPAEPK-SAPPPAEPKSAEPKPAEPKSGPPAEPKSAEPKPAEPK 63
QY 174 VAAEPAAEPVGEPAEPSP-----EPATAPPAEPSPSPS 213
DB 64 KPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPK 108

RESULT 9

translational initiation factor IF-2 - Stigmatella aurantiaca
C:Species: Stigmatella aurantiaca
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Feb-2001
A:Accession: T43226
R:Bremund, L.; Laalami, S.; Derijard, B.; Cennatempo, Y.
J. Bacteriol. 179, 2348-2355, 1997
A:Title: Translational initiation factor IF2 of the myxobacterium Stigmatella aurantiaca:
A:Reference number: Z22352; MUID:97234648; PMID:9079922
A:Accession: T43226
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1054 <BRE>
A:Cross-references: EMBL:X87940; PIDN:CAA61162.1
A:Experimental source: strain DM4
C:Genetics:
A:Gene: infB
C:Superfamily: translation initiation factor IF-2; translation elongation factor Tu homc
C:Keywords: nucleotide binding; P-loop; protein biosynthesis
F:555-664/Domain: translation elongation factor Tu homology <ETU>
F:561-568/Region: nucleotide-binding motif A (P-loop)

Query Match 14.9%; Score 175; DB 2; Length 1054;
Best Local Similarity 38.3%; Pred. No. 0.0013;
Matches 49; Conservative 8; Mismatches 51; Indels 20; Gaps 4;

QY 96 ATQANVPLPRGPFVPSRFFSAAAAPAPPIA--EPAAAAPLTAT--PVAEPAA 151
DB 98 ASDVSSPSP-----VHEAGAEAAASERVPEAAAVQEPVAEAPRAAASEPAAAPKAT 152
QY 152 APVAEPAAEPVGAEPAAEPVAAEPAAEPVGEPAEPSP-----APAT 200
DB 153 APVAEPVVEAPKAAAPVAEPVVEAPKTEAPVAAAPVIAENPTTPARTTEVPTSGRAAS 212
QY 201 AKPAPEP 208
DB 213 CRGAAPLP 220

RESULT 10

QOBER
BFLF1 protein - human herpesvirus 4 (strain B95-8)
C:Species: human herpesvirus 4, Epstein-Barr virus
C>Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 16-Jul-1999


```

Db      16 GVDSDTAVNMDSPT-ASDGPRIEVSAGAPVSEHAKRPLEHQALFQNSPISSTTAAEA 74
Qy      71 GLPSTPWLITSPGFPYVYHIRGFLATQ-----LNVPPLPRGFPVPP 114
Db      75 KVPSTLEREGSGPTQETVHIKPAVPAESGTDSSKADPSATAVIQTPEEVEGVPTMPT 134
Qy      115 SRFPSAAAAAPPAAPPAAP-AAAAPLTATPVAA-----BPAGAPVAAEPAAAPVGAAPA 169
Db      135 D-LEPPASEDGGPDVRAEPDGGTAP--ATPAESENREPAAAA--AAEPAAAP--AAEPA 186
Qy      170 AE--APVAAEPAAEA-----PVGVEPAAEPSAPAPATAPAPAPESP 212
Db      187 AEPAAPPAEPAAEAVPDTAEASASGAVPDTQEBPAAAAASATPAPAPARAAP 239

```

RESULT 14 G87675

ATPase-related protein (imported) - Caulobacter crescentus
 C.Species: Caulobacter crescentus
 C.Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C.Accession: G87675
 R.Nielsen, W.C.; Feildjlym, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon
 n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A.Title: Complete Genome Sequence of Caulobacter crescentus.
 A.Reference number: A87249; MUID:21173698; PMID:11259647
 A.Accession: G87675
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-438 <STO>
 A.Cross-references: GB:AE005673; NID:g13425157; PIDN:AAK25403.1; GSPDB:GN00148
 A.Genetic8:
 A.Gene: CC3441

```

Query Match      14.6%; Score 171.5; DB 2; Length 438;
Best Local Similarity 33.3%; Pred. No. 0.00094;
Matches 68; Conservative 14; Mismatches 59; Indels 63; Gaps 11;

Qy      64 GNTYDTGLPSYP-----WLTSPGF-----PYVYHIRGFLATQNLVPPPLPRGFPVPP 114
Db      219 GRKVASADIKIYNNNAHLIDEPNTMKLIDVIAWLKGTQAPAKVBPAPA----- 270
Qy      115 SRFPSAAAAAP-APPIAEPAAAPLTATPV-AAEPAA-----GAPVAAEPA-----A 160
Db      271 ----PAADAPAPAPAKAPEPAAP--EPVKAEPAPSPKAKAPKAPKADAKPKATA 323
Qy      161 EAPVG-----APPAEAPV--AAEPAAEPVGEPAEESPAPF----- 197
Db      324 KAPVAKKAAAPKAKAAPAAEAPKSAAPKAPKAPKAAAPKAAAPKAEKPKTAAKAPAAE 383
Qy      198 --PATAPAPAPSPSPLEQANQ 219
Db      384 TAPAKKAPAPKAAKPAKATATK 407

```

RESULT 15 T17547

proline-rich protein A57R - Chlorella virus PBCV-1
 C.Species: Chlorella virus PBCV-1
 C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C.Accession: T17547
 R.Graves, M.V.; Van Etten, J.L.
 submitted to the EMBL Data Library, May 1999
 A.Reference number: Z18806
 A.Accession: T17547
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-544 <GRA>
 A.Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96425.1
 A.Experimental source: specific host Chlorella strain NC64A
 C.Genetic8:

```

A.Gene: A57R
Query Match      14.5%; Score 171; DB 2; Length 544;
Best Local Similarity 27.8%; Pred. No. 0.0012;
Matches 58; Conservative 19; Mismatches 106; Indels 26; Gaps 4;

Qy      2 KLLWACIVCAFAFKRRF---PFIQEDNDGHPHPSLNIPIGIRLPPPLYRPVNT 58
Db      4 RVIFFLITVTTAFARPEPFTTNOIVTQTTVNSPKPSPKPPSPKPPSPKPPSP 63
Qy      59 VPSYGNVTYDTGLPSYPWLTSPGFPYVYHIRGFLATQNLVPPPLPRGFPVPPSRFP 118
Db      64 KPSPKPPSPKPPSPKPP-----SP-----KPPSPKPPSP-KPPSPKP 100
Qy      119 SAAAPPAAPPIAEPAAAPLTATPVAAEPAAAGAPVAAEAPVGAEPAAEAPVAAEP 178
Db      101 PSPKPPSPKPPSPKPPSPKPPSPKPPSPKPPSPKPPAPAPAPAPAPAPAPAP 160
Qy      179 AAAPVGEPAEESPAPAPATAPAPAE 207
Db      161 PAPBPALBPAPAPAPAPAPAPAPAPAE 189

```

Search completed: January 15, 2004, 15:14:44
 Job time : 23 secs

THIS PAGE BLANK (USPTO)

Db 181 EAPVGEPAABESPAPATAPAAEPSPSLEQANQ 219

RESULT 2
US-09-923-236-2
Sequence 2, Application US/09922469
Patent No. US20020090677A1
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Adler, David A.
TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
FILE REFERENCE: 97-71
CURRENT APPLICATION NUMBER: US/09/923,236
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US 60/124,820
PRIOR FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 2
LENGTH: 219
TYPE: PRT
ORGANISM: Homo sapiens
US-09-923-236-2

Query Match 100.0%; Score 1178; DB 9; Length 219;
Best Local Similarity 100.0%; Pred. No. 2,6e-77;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLMACIVCAFAKRRFPPIGEDNDGPHLPSLNIPIYGINLPPPLYRRPVNTVP 60
1 MLLLMACIVCAFAKRRFPPIGEDNDGPHLPSLNIPIYGINLPPPLYRRPVNTVP 60
Db 1 MLLLMACIVCAFAKRRFPPIGEDNDGPHLPSLNIPIYGINLPPPLYRRPVNTVP 60

QY 61 SYPGNTYTDGTPSYWILTSRQFPYVYHIRGFPLATQLVNPPPLPRGPPFVPSRFFSA 120
61 SYPGNTYTDGTPSYWILTSRQFPYVYHIRGFPLATQLVNPPPLPRGPPFVPSRFFSA 120
Db 61 SYPGNTYTDGTPSYWILTSRQFPYVYHIRGFPLATQLVNPPPLPRGPPFVPSRFFSA 120

QY 121 AAAPAPPIAAEPAAAPLTATPVAAEPAAAGAPVAAEPAAEPVGAEPAAEPVAAEPAA 180
121 AAAPAPPIAAEPAAAPLTATPVAAEPAAAGAPVAAEPAAEPVGAEPAAEPVAAEPAA 180
Db 121 AAAPAPPIAAEPAAAPLTATPVAAEPAAAGAPVAAEPAAEPVGAEPAAEPVAAEPAA 180

QY 181 EAPVGEPAABESPAPATAPAAEPSPSLEQANQ 219
181 EAPVGEPAABESPAPATAPAAEPSPSLEQANQ 219
Db 181 EAPVGEPAABESPAPATAPAAEPSPSLEQANQ 219

RESULT 3
US-09-922-469-2
Sequence 2, Application US/09922469
Patent No. US20020173027A1
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Adler, David A.
TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
FILE REFERENCE: 97-71
CURRENT APPLICATION NUMBER: US/09/922,469
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US 60/124,820
PRIOR FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 2
LENGTH: 219
TYPE: PRT
ORGANISM: Homo sapiens
US-09-922-469-2

Query Match 100.0%; Score 1178; DB 10; Length 219;
Best Local Similarity 100.0%; Pred. No. 2,6e-77;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLMACIVCAFAKRRFPPIGEDNDGPHLPSLNIPIYGINLPPPLYRRPVNTVP 60
1 MLLLMACIVCAFAKRRFPPIGEDNDGPHLPSLNIPIYGINLPPPLYRRPVNTVP 60
Db 1 MLLLMACIVCAFAKRRFPPIGEDNDGPHLPSLNIPIYGINLPPPLYRRPVNTVP 60

Db 181 EAPVGEPAABESPAPATAPAAEPSPSLEQANQ 219

RESULT 4
US-09-746-783-106
Sequence 106, Application US/09746783
Publication No. US20030044935A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavallee, Edward R.
Racie, Lisa A.
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.
Howes, Steven H.
Fechtel, Kim
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/746,783
FILING DATE: 21-Dec-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Malaicinic, Debra J.
REGISTRATION NUMBER: 46,931
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 219 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 106:
US-09-746-783-106

Query Match 99.5%; Score 1172; DB 11; Length 219;
Best Local Similarity 99.1%; Pred. No. 7e-77;
Matches 217; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLMACIVCAFAKRRFPPIGEDNDGPHLPSLNIPIYGINLPPPLYRRPVNTVP 60
1 MLLLMACIVCAFAKRRFPPIGEDNDGPHLPSLNIPIYGINLPPPLYRRPVNTVP 60
Db 1 MLLLMACIVCAFAKRRFPPIGEDNDGPHLPSLNIPIYGINLPPPLYRRPVNTVP 60

QY 61 SYPGNTYTDGTPSYWILTSRQFPYVYHIRGFPLATQLVNPPPLPRGPPFVPSRFFSA 120
61 SYPGNTYTDGTPSYWILTSRQFPYVYHIRGFPLATQLVNPPPLPRGPPFVPSRFFSA 120
Db 61 SYPGNTYTDGTPSYWILTSRQFPYVYHIRGFPLATQLVNPPPLPRGPPFVPSRFFSA 120

Db 61 SYPGNTYDTGLPSYFWILTSRGFYVYHINGFPLATQUNVPLPRGPFVPSRFSFA 120
 QY 121 AAAPAPPIAAEPAAAAELTATPVAAEPAGAPVAAEPAAEPVAAEPVAAEPAA 180
 Db 121 AAAPAPPIAAEPAAAAELTATPVAAEPAGAPVAAEPAAEPVAAEPVAAEPAA 180
 QY 181 EAPVGEPAEAEPSPAEPATKAPAPBPHPSPLEQANQ 219
 Db 181 EAPVGEPAEAEPSPAEPATKAPAPBPHPSPLEQANQ 219

RESULT 5

US-10-231-417-192
 ; Sequence 192, Application US/10231417
 ; Publication No. US20030176681A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Feng et al.
 ; TITLE OF INVENTION: 148 Human Secreted Proteins
 ; FILE REFERENCE: P2019P1
 ; CURRENT APPLICATION NUMBER: US/10/231,417
 ; CURRENT FILING DATE: 2002-08-30
 ; PRIOR APPLICATION NUMBER: US/09/296,622
 ; PRIOR FILING DATE: 1999-04-23
 ; NUMBER OF SEQ ID NOS: 619
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 192
 ; LENGTH: 221
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1159)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (221)
 ; OTHER INFORMATION: Xaa equals stop translation
 ; US-10-231-417-192

Query Match 94.5%; Score 1113.5; DB 12; Length 221;
 Best Local Similarity 95.0%; Pred. No. 1.1e-72;
 Matches 209; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

QY 1 MGLIMACTVCVAFARKRRFPETGGDNDGHPHLSINIPYGINLPPPLYRPVNTVP 60
 Db 1 MGLIMACTVCVAFARKRRFPETGGDNDGHPHLSINIPYGINLPPPLYRPVNTVP 60
 QY 61 SYPGNTYDTGLPSYFWILTSRGFYVYHINGFPLATQUNVPLPRGPFVPSRFSFA 120
 Db 61 SYPGNTYDTGLPSYFWILTSRGFYVYHINGFPLATQUNVPLPRGPFVPSRFSFA 120
 QY 121 AAAPAPPIAAEPAAAAELTATPVAAEPAGAPVAAEPAAEPVAAEPVAAEPAA 179
 Db 121 AAAPAPPIAAEPAAAAELTATPVAAEPAGAPVAAEPAAEPVAAEPVAAEPAA 180
 QY 180 AEPVGEPAEAEPSPAEPATKAPAPBPHPSPLEQANQ 219
 Db 181 AEPVGEPAEAEPSPAEPATKAPAPBPHPSPLEQANQ 220

RESULT 6

US-10-316-253-95
 ; Sequence 95, Application US/10316253
 ; Publication No. US20030162706A1
 ; GENERAL INFORMATION:
 ; APPLICANT: The Procter & Gamble Company
 ; APPLICANT: Peters, Kevin
 ; APPLICANT: Thompson, Larry
 ; APPLICANT: Wang, Feng
 ; APPLICANT: Greis, Kenneth
 ; TITLE OF INVENTION: Angiogenesis Modulating Proteins
 ; FILE REFERENCE: 8865M

; CURRENT APPLICATION NUMBER: US/10/316,253
 ; CURRENT FILING DATE: 2002-12-10
 ; PRIOR APPLICATION NUMBER: US 60/355,295
 ; PRIOR FILING DATE: 2002-02-08
 ; NUMBER OF SEQ ID NOS: 308
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 95
 ; LENGTH: 2657
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-10-316-253-95

Query Match 15.4%; Score 181; DB 12; Length 2657;
 Best Local Similarity 29.1%; Pred. No. 0.00012;
 Matches 57; Conservative 22; Mismatches 61; Indels 56; Gaps 8;

QY 40 IPYGINLPP-----LYRPVNTVPSYNGNTYDTGLPSYFWILTSRGFYVYHINGF 93
 Db 2239 VKFGHKQINPHTNSSLTSKVTTIK--PNSSKPTAI-----VNLTP----- 2280
 QY 94 PLATQUNVPLPRGPFV-----PSRFSAAAAPAP-----PIAEP 133
 Db 2281 -----AKPAPAPAPAPQPVLAKEPPAKQAPAPKPSAKLVPPQPVHVQAPAPQT 2333
 QY 134 AAAPPLTATPVAAEPAGAPVAAEPAAEA-PVGAEPAAEPVAAEPAAEPVGEPAE 192
 Db 2334 ASVPPAPKAPAPQAPAAKVPAPKAPVAPQAPAPAPAPAPAAKVPAPKAVPA-----OPAAQ 2389
 QY 193 PPAEPATKAPAPAP 208
 Db 2390 PMPAPVLTAKSAAVKP 2405

RESULT 7

US-10-264-049-2409
 ; Sequence 2409, Application US/10264049
 ; Publication No. US20040005579A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Biree et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PA133P1
 ; CURRENT APPLICATION NUMBER: US/10/264,049
 ; CURRENT FILING DATE: 2002-10-04
 ; PRIOR APPLICATION NUMBER: PCT/US01/18569
 ; PRIOR FILING DATE: 2001-06-07
 ; PRIOR APPLICATION NUMBER: US 60/209,467
 ; PRIOR FILING DATE: 2000-06-07
 ; NUMBER OF SEQ ID NOS: 4360
 ; SOFTWARE: PatentIn Ver. 3.1
 ; SEQ ID NO 2409
 ; LENGTH: 903
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: MISC_FEATURE
 ; LOCATION: (110)
 ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 ; FEATURE:
 ; NAME/KEY: MISC_FEATURE
 ; LOCATION: (15)
 ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 ; FEATURE:
 ; NAME/KEY: MISC_FEATURE
 ; LOCATION: (16)
 ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 ; FEATURE:
 ; NAME/KEY: MISC_FEATURE
 ; LOCATION: (795)
 ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
 ; US-10-264-049-2409

Query Match 14.9%; Score 175.5; DB 12; Length 903;
 Best Local Similarity 32.3%; Pred. No. 9.9e-05;

TYPE: PRT
ORGANISM: Homo sapiens
US-10-023-523-43

Query Match
Best Local Similarity 14.7%; Score 173; DB 14; Length 538;
36.8%; Pred. No. 8.9e-05;
Matches 49; Conservative 7; Mismatches 47; Indels 30; Gaps 5;

QY 102 PPLPRGPFVPPSRFFSAAAPLTAAPPAAPLTAATVAAEPAAAGVAAEPAE 161
DB 111 PPRARGPAP-----AAAAAAPPPTAPPPPPAPVAAAAPAAAAAATAPSP 162
QY 162 APVGAEPAAE-APVAA-EPAEAPVGVPE-----AABEP-----SPAPA 199
DB 163 GAAGGPRAGRAAPLAPPAAPPAAPVAPAGPRAPPAVAAEPPLPPPPQPPAPQ 222
QY 200 TAKPAAPEPHPS 212
DB 223 QQQPPPPQPP 235

RESULT 11

US-10-029-386-33686
Sequence 33686, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
FILE REFERENCE: AEWITCA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Seqmax Sequence Listing Engine ver. 1.1
SEQ ID NO: 33686
LENGTH: 980
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO ALI37853.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.84
OTHER INFORMATION: SWISSPROT HIT: Q02910, EVALUE 2.00e-03
US-10-029-386-33686

Query Match
Best Local Similarity 14.7%; Score 173; DB 12; Length 980;
31.1%; Pred. No. 0.00016;
Matches 68; Conservative 13; Mismatches 112; Indels 26; Gaps 9;

QY 6 WACIVCAFAARRRRPFIGEDNDGDHPLHPSLNT-----PYGIRNLPPPLYRPNVT 58
DB 531 WAAIAITV-----PITBEDGTPBG-PVTPTATVAAPEPDTAARVAVSTPREPASPA 582
QY 59 VSYSGNTYTDGLSGYPILTSRGPYVYHNRGFLATQLVNPPRPGPFVP-PSRF 117
DB 563 VTPPEPTSPAAAVPT-PEEPTSPAAAVPPPEPTSPAAAVPTPEEPTSPAAAVPTPEEP 641
QY 118 FSAAPAAPAPLTAAPPAAPLTAATVAAEPAAAGVAAEPAE 174
DB 642 TSPAAAVPTPEEPTSPAAAVPTPEEPTSPAAAVPTPEEPTSPAAAVPTPEEPA-SBA 697
QY 175 AAEPAEAPVGAEPAAEPPSPAEPAATAPKPAEPPHPS 213
DB 698 AAVPTPEEP--ASPAAPVPTPEEPAPPAVAVPTPEESAS 724

RESULT 12
US-10-108-260A-3849
Sequence 3849, Application US/10108260A
Publication No. US20040005560A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE

TITLE OF INVENTION: NO. US20040005560A1 full length cDNA

FILE REFERENCE: H1-A0106
CURRENT APPLICATION NUMBER: US/10/108,260A
CURRENT FILING DATE: 2002-03-27
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO: 3849
LENGTH: 693
TYPE: PRT
ORGANISM: Homo sapiens
US-10-108-260A-3849

Query Match
Best Local Similarity 14.6%; Score 171.5; DB 12; Length 693;
31.6%; Pred. No. 0.00015;
Matches 62; Conservative 17; Mismatches 78; Indels 39; Gaps 11;

QY 33 PLHPSLNPYGIRNLPPPLYRPNVTVPSYQNTYTDGLSPYILTSRGPYVYHNRG 92
DB 253 PVSPEATP-----SQAPFLAAPPLQVPPSPASP-----PMSPEATPPQAP----- 297
QY 93 PPLAT-QLVNPPPLPRGPFVPPSRFFSAAAP-----AAPLAEPAAPLTAATV 145
DB 298 -PLAAPPLQVPPSP 351
QY 146 AEPAGAVAAEPAAEAPVGAEPAAE-APVAAEPAAEAPVGAEPAAEPPSPAPATK 202
DB 352 KPPQAPPALATPPLQA-LPSPPASFGQAPPS-PSASLP--MSPLATPPQAPPLVLA 406
QY 203 P---AAPEHPSPSLE 215
DB 407 PLQVPPSP 422

RESULT 13

US-10-084-843-142
Sequence 142, Application US/10084843
Publication No. US20030143243A1

GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Iwardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
AND DIAGNOSIS OF TUBERCULOSIS

NUMBER OF SEQUENCES: 355

CORRESPONDENCE ADDRESS:
ADDRESS: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle

STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/084,843
FILING DATE: 25-Feb-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/072,967

FILING DATE: 05-MAY-1998

ATTORNEY/AGENT INFORMATION:

NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 142:
: SEQUENCE CHARACTERISTICS:
:     LENGTH: 267 amino acids
:     TYPE: amino acid
:     STRANDEDNESS: single
:     TOPOLOGY: linear
: MOLECULE TYPE: peptide
: SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-10-084-843-142

Query Match      14.4%; Score 170; DB 12; Length 267;
Best Local Similarity 28.0%; Pred. No. 7, 2e-05;
Matches 56; Conservative 14; Mismatches 94; Indels 36; Gaps 11.

Cy      21 PFIGEDDDNDGHPPLPSINIPY--GIRULPPLYYRPVN-TVPSYPGNTYTDGLPSYWM 77
Db      79 PRFGTSPGVGGPRPAPPAPEAPVAVPVI---PVIILIPPPG---WQGNPITLP- 131
Cy      78 ILTSGGFVYVTHIRGFPLATQLNVPPLPPRGPFVPPSGRFSMAAALAAPAPPIAAEPAANA 137
Db      132 --TAE--PTT-----PVTTATPTPTPTPTPTPTPT-----PTTP--TTPTVTP 170
Cy      138 PLT--ATPVAAPPAAGAVPAAPAAEAPVAGAPAAEAPVAAPAAEAPVAGVAPAAEES- 194
Db      171 PTTPTPTVTPTPTPTTVAPTTVAPTTVAPTTVAPTTVAATATPTTVAAPQTQPPQTQ 230
Cy      195 --PAEPATAKPAAPPEPSP 212
Db      231 QMPTQQTVAAPQTVAAPQP 250

```

RESULT 14
 US-10-193-002-137
 Sequence 137, Application US/10193002
 Publication No. US20030135026A1
 GENERAL INFORMATION:
 APPLICANT: Reed, Steven G.
 Skeiky, Yasir A.W.
 Dillon, Davin C.
 Campos-Neto, Antonia
 Houghton, Raymond
 Vedvick, Thomas S.
 Twardzik, Daniel R.
 Iodes, Michael J.
 Hendrickson, Ronald C.
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
 TUBERCULOSIS
 NUMBER OF SEQUENCES: 350
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED AND BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104-7092
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/193, 002
 FILING DATE: 10-Jul-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/072,596
 FILING DATE: 05-MAY-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Maki, David J.
 REGISTRATION NUMBER: 31,392

```

REFERENCE/DOCKET NUMBER: 210121.417C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 137:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 137:
us-10-193-002-137

Query Match          14.4%; Score 170; DB 12; Length 267;
Best Local Similarity 28.0%; Pred. No. 7,2e-05;
Matches 56; Conservative 14; Mismatches 94; Indels 36; Gaps 11

QY      21 PFIGEDNDGDGHPHPISINIPY--GIRNLPPPLLYRPN-TVPSYEGNTYDTGLPSIPW 77
DB      79 PRPGTSPGVGCTGPSAPSEAPVAGVAVPAPI--PVPILIPPEFG--NQGMPTIP- 131
QY      78 ILTSRGPFYVYHNGFPLATQNLNVPLEBRGFFPFVPSRFSAAAPARAAPRTIAEPAA 137
DB      132 --TAP--PTT-----PVTISATTPPTTPTTPTTPTT-----PTTPP--TTPTVTP 170
QY      138 PLT--ATPVAAPAGAFVAAPAAAPAAAPVGAAPAAAPVAAPAAAPVGAAPAAEES- 194
DB      171 PTPPTPTPTPTPTPTTVAAPTVAAPTVAAPTVAAPTVAAPTVAAPTVAAPTQPTQPTQPTQ 230
QY      195 --PAEPATAKPAEPHPSP 212
DB      231 QMPTQOQTVAAPQVAPAPQP 250

```

```

RESULT 15
US-09-906-514-4
; Sequence 4, Application US/09906514
; Patent No. US20020170085A1
; GENERAL INFORMATION:
; APPLICANT: Kaeppler, Shawn
; APPLICANT: Springer, Nathan
; APPLICANT: Phillips, Ronald
; TITLE OR INVENTION: Methyl Cpg Binding Domain Nucleic Acids from Maize
; FILE REFERENCE: Methyldinding
; CURRENT APPLICATION NUMBER: US/09/906,514
; CURRENT FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Zea mays
US-09-906-514-4

Query Match          14.4%; Score 170; DB 10; Length 428;
Best Local Similarity 33.3%; Pred. No. 0.00012;
Matches 74; Conservative 13; Mismatches 89; Indels 46; Gaps 11.

25 EDDNDGHLPLPSLNIPIGIRNLPEPLLYRPVNTVPSYGTATYDTGLPSYFWITTSQGF 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
183 EKKEDAKPDEPBAVAPAVSN--PTENSASAPAPBAVPAVPVETESVAPPAVLAA- 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25 PYVHIRGFPLATQLNVLPPLPRGCPFVPSRFFSAAA-----PAAPF----- 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
240 PETK----PDAKPAVAPAPAPENKPDAP-----AAAAVAVDPTKSAEPAAAPDPTKS 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
130 AAEPPAAAPLTATPVAEAPAGAPVAAEPAAAPVGAEPAAEAPV-----AAEPAAEAPV 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
291 VAEPPAAAPVETKLVAESADA--VAAPAPETKSDAEPA--APVETKPVAESADAIVA 347
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
185 GVEPAAEESPAPATAKPAPE-----PHPSPELEQA 217
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Tue Jan 20 17:43:12 2004

us-09-923-236-2.rapb

Page 7

Db 348 A--PAPEKSDAEPAAADPAPEIKSDAAAADPAFGKADA 387

Search completed: January 15, 2004, 15:19:47
Job time : 34 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:11:59 ; Search time 21 Seconds

(without alignments)
441.242 Million cell updates/sec

Title: US-09-923-236-2

Perfect score: 1178

Sequence: 1 MKLLIMACTIVCAFAKRRF.....TAKAAPRHPSLEQANQ 219

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1178	100.0	219	4	US-09-527-345-2
2	179	15.2	207	1	US-07-836-642-2
3	179	15.2	207	1	US-08-169-563-2
4	179	15.2	207	1	US-08-403-379A-6
5	179	15.2	207	2	US-08-557-309B-50
6	179	15.2	207	2	US-08-929-414-6
7	179	15.2	207	3	US-08-834-306-50
8	179	15.2	207	3	US-08-993-674A-50
9	179	15.2	207	4	US-09-256-976-50
10	170	14.4	267	3	US-08-818-112-142
11	170	14.4	267	4	US-08-818-111-137
12	170	14.4	267	4	US-09-056-556-142
13	170	14.4	267	4	US-09-072-596-137
14	164.5	13.9	316	4	US-09-252-991A-32957
15	164.5	13.9	316	4	US-08-217-327-4
16	154	13.1	805	3	US-09-103-429A-4
17	153	13.0	786	3	US-09-103-429A-3
18	150	12.7	399	4	US-09-252-991A-22853
19	148	12.6	2972	4	US-09-579-181-2
20	148	12.6	3118	4	US-09-579-181-1
21	145.5	12.4	1565	6	5352450-2
22	144	12.2	195	4	US-09-252-991A-20967
23	142.5	12.1	515	4	US-09-252-991A-27942
24	141	12.0	369	4	US-09-252-991A-25394
25	139.5	11.8	941	4	US-07-757-022B-14
26	139.5	11.8	1022	4	US-07-757-022B-84
27	139.5	11.8	1038	4	US-07-757-022B-74

28	139.5	11.8	1049	4	US-07-757-022B-58	Sequence 58, Appl
29	139.5	11.8	1140	4	US-07-757-022B-104	Sequence 104, App
30	139.5	11.8	1270	4	US-07-757-022B-44	Sequence 44, Appl
31	139.5	11.8	1311	4	US-07-757-022B-42	Sequence 42, Appl
32	139.5	11.8	1313	4	US-07-757-022B-142	Sequence 142, App
33	139.5	11.8	1314	4	US-07-757-022B-50	Sequence 50, Appl
34	139.5	11.8	1320	4	US-07-757-022B-46	Sequence 46, Appl
35	139.5	11.8	1320	4	US-07-757-022B-60	Sequence 60, Appl
36	139.5	11.8	1354	4	US-07-757-022B-48	Sequence 48, Appl
37	139.5	11.8	1361	4	US-07-757-022B-40	Sequence 40, Appl
38	139.5	11.8	1363	4	US-07-757-022B-52	Sequence 52, Appl
39	139.5	11.8	1404	4	US-07-757-022B-2	Sequence 2, Appl1
40	139.5	11.8	1404	4	US-07-757-022B-62	Sequence 62, Appl1
41	138	11.7	180	6	5273901-7	Patent No. 5482709
42	138	11.7	180	6	5482709-6	Patent No. 5482709
43	138	11.7	304	4	US-09-252-991A-23116	Sequence 23116, A
44	136.5	11.6	504	3	US-09-219-849-3	Sequence 3, Appl1
45	136.5	11.6	561	1	US-08-642-255-52	Sequence 52, Appl

ALIGNMENTS

```
RESULT 1
US-09-527-345-2
Sequence 2, Application US/09527345
Patent No. 6331413
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
TITLE OF INVENTION: SECRETED SALIVARY ZSIC63 POLYPEPTIDE
FILE REFERENCE: 97-71
CURRENT APPLICATION NUMBER: US/09/527,345
CURRENT FILING DATE: 1999-03-17
PRIOR APPLICATION NUMBER: US 60/124,820
PRIOR FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 219
TYPE: PRT
ORGANISM: Homo sapiens
US-09-527-345-2

Query Match      100.0%; Score 1178; DB 4; Length 219;
Best Local Similarity 100.0%; Pred. No. 3.3e-93;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKLLIMACTIVCAFAKRRFPFIEDDDHPLHPSLNIPYGRNLPPLYYRPNVTP 60
      |||
DB      1 MKLLIMACTIVCAFAKRRFPFIEDDDHPLHPSLNIPYGRNLPPLYYRPNVTP 60
      |||

QY      61 SYPGNTYDTGLPSYFWILTSFGFPYVYHNGFPLATOLNVPPLPRGFPVPSRFSQA 120
      |||
DB      61 SYPGNTYDTGLPSYFWILTSFGFPYVYHNGFPLATOLNVPPLPRGFPVPSRFSQA 120
      |||

QY      121 AAAPAPPIAEPAAALUTITPVAEPAAGAPVAEPAEPAEPAEPAEPAEPA 180
      |||
DB      121 AAAPAPPIAEPAAALUTITPVAEPAAGAPVAEPAEPAEPAEPAEPAEPA 180
      |||

QY      181 EAPGVPEAPAEPPAPAPATKAPAPRHPSPSLEQANQ 219
      |||
DB      181 EAPGVPEAPAEPPAPAPATKAPAPRHPSPSLEQANQ 219
      |||

RESULT 2
US-07-836-642-2
Sequence 2, Application US/07836642
Patent No. 5304371
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: Peptide for Diagnosing and Immunizing
      Against T. cruzi Infection
```

```

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSER: Immunex Corporation
STREET: 51 University street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/836,642
FILING DATE: 14-FEB-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: REED
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 207 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-836-642-2

Query Match      15.2%; Score 179; DB 1; Length 207;
Best Local Similarity 41.0%; Pred. No. 4.1e-08;
Matches 43; Conservative 11; Mismatches 45; Indels 6; Gaps 2;

QY 114 PSRFFSAAAPAPPIAEPAAAPLTATPVAAEPAAAGPVAAEPAAEPVGAEPAAEP 173
DB 5 PAEPKSAEPKPAEPK-SAEPPKPAEPKSAEPKPAEPKSAEPKPAEPKPAEPKPAEP 63
QY 174 VAAEPAAEPVGVPEAAEPSP-----AEPATKPAEPDPHPS 213
DB 64 KPAPKSAEPKPAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEP 108

RESULT 3
US-08-169-563-2
Sequence 2, Application US/08169563
Patent No. 5413912
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: Peptide for Diagnosing and Immunizing
TITLE OF INVENTION: Against T. cruzi Infection
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSER: Immunex Corporation
STREET: 51 University street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple Operating System 7.1
SOFTWARE: Microsoft Word For Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/169,563
FILING DATE: 17-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/836,642
FILING DATE: 14-FEB-1992
CLASSIFICATION: 435
```

```

ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: REED-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 207 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-169-563-2

Query Match      15.2%; Score 179; DB 1; Length 207;
Best Local Similarity 41.0%; Pred. No. 4.1e-08;
Matches 43; Conservative 11; Mismatches 45; Indels 6; Gaps 2;

QY 114 PSRFFSAAAPAPPIAEPAAAPLTATPVAAEPAAAGPVAAEPAAEPVGAEPAAEP 173
DB 5 PAEPKSAEPKPAEPK-SAEPPKPAEPKSAEPKPAEPKSAEPKPAEPKPAEPKPAEP 63
QY 174 VAAEPAAEPVGVPEAAEPSP-----AEPATKPAEPDPHPS 213
DB 64 KPAPKSAEPKPAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEPKSAEP 108

RESULT 4
US-08-403-379A-6
Sequence 6, Application US/08403379A
Patent No. 5756662
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION
TITLE OF INVENTION: OF T. CRUZI INFECTION
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSER: SED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,379A
FILING DATE: 14-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Naki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.406
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 207 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-403-379A-6

Query Match      15.2%; Score 179; DB 1; Length 207;
Best Local Similarity 41.0%; Pred. No. 4.1e-08;
Matches 43; Conservative 11; Mismatches 45; Indels 6; Gaps 2;

QY 114 PSRFFSAAAPAPPIAEPAAAPLTATPVAAEPAAAGPVAAEPAAEPVGAEPAAEP 173
```

Db 5 PAEPKSAEPKPAEPK-SAEPPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP 63
Qy 174 VAAPPAAPVGVPEPAEPPSP-----AEPTAKPAEPHPSPS 213
Db 64 KPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP 108

RESULT 5
US-08-557-309B-50
; Sequence 50, Application US/08557309B
; Patent No. 5916572
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,309B
FILING DATE: 14-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.422
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 207 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

US-08-557-309B-50

Query Match 15.2%; Score 179; DB 2; Length 207;
Best Local Similarity 41.0%; Pred. No. 4,1e-08;
Matches 43; Conservative 11; Mismatches 45; Indels 6; Gaps 2;

Qy 114 PSRFSAAAAPAPPIAAEPAAPLTATPVAAEPKAGAPVAAEPKAAAPGAEPPAAP 173
Db 5 PAEPKSAEPKPAEPK-SAEPPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP 63

Qy 174 VAAPPAAPVGVPEPAEPPSP-----AEPTAKPAEPHPSPS 213
Db 64 KPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP 108

RESULT 6
US-08-929-414-6
; Sequence 6, Application US/08929414
; Patent No. 5942403
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Houghton, Raymond
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION
; TITLE OF INVENTION: OF T. CRUZI INFECTION
; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/929,414
FILING DATE: 15-SEP-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.406C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 207 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-929-414-6

Query Match 15.2%; Score 179; DB 2; Length 207;
Best Local Similarity 41.0%; Pred. No. 4,1e-08;
Matches 43; Conservative 11; Mismatches 45; Indels 6; Gaps 2;

Qy 114 PSRFSAAAAPAPPIAAEPAAPLTATPVAAEPKAGAPVAAEPKAAAPGAEPPAAP 173
Db 5 PAEPKSAEPKPAEPK-SAEPPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP 63

Qy 174 VAAPPAAPVGVPEPAEPPSP-----AEPTAKPAEPHPSPS 213
Db 64 KPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP 108

RESULT 7
US-08-834-306-50
; Sequence 50, Application US/08834306
; Patent No. 6054135
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/834,306
FILING DATE: 15-APR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.

```

?      REGISTRATION NUMBER:   31,392
?      REFERENCE/DOCKET NUMBER: 210121.422C1
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: (206) 622-4800
?      TELEFAX: (206) 682-6031
?      INFORMATION FOR SEQ ID NO: 50:
?      SEQUENCE CHARACTERISTICS:
?          LENGTH: 207 amino acids
?          TYPE: amino acid
?          STRANDEDNESS:
?              TOPOLOGY: linear
JS-08-834-306-50

```

Query Match	15.2%;	Score 179;	DB 3;	Length 207;
Best Local Similarity	41.0%;	Pred. No. 4.1e-08;		
Matches 43;	Conservative 11;	Mismatches 45;	Indels 6;	Gaps 2.

Qy	Db	Qy	Db
114	5	174	64
PSSEFFSAAAPAAPPLTAAEAPAAAPLTTATPVAAEPAGAPVAAEAPAAAPVCAEAPAAAP	PAEPKSAEPKPAEPK-SAEPPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP	VAAEPAAEAIVGVEPAAEESP-----AAPATKAPAAEPHPHSPS	KPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP
174	63	213	108
PSSEFFSAAAPAAPPLTAAEAPAAAPLTTATPVAAEPAGAPVAAEAPAAAPVCAEAPAAAP	PAEPKSAEPKPAEPK-SAEPPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP	VAAEPAAEAIVGVEPAAEESP-----AAPATKAPAAEPHPHSPS	KPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEPKPAEPKSAEP

RESULT 8
US-08-993-674A-50
; Sequence 50, Application US/08993674A

```

1  GENERAL INFORMATION:
2  APPLICANT: Reed, Steven G.
3  APPLICANT: Skelky, Yasir A.W.
4  APPLICANT: Iodes, Michael J.
5  APPLICANT: Houghton, Raymond L.
6  APPLICANT: Smith, John M.
7  APPLICANT: McNeill, Patricia D.
8  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND
9  NUMBER OF SEQUENCES: 81
10

```

Query Match	15.2%;	Score 179;	DB 3;	Length 207;
Best Local Similarity	41.0%;	Pred. No. 4.1e-08;		
Matches	43;	Conservative	11;	Mismatches 45;
			Indels	6;
			Gaps	2

[illegible]

RESULT 9
US-09-256-976-50
; Sequence 50, Application US/09256976

```

1  APPLICANT: Reed, Steven G.
2  APPLICANT: Skeiky, Yasar A.W.
3  APPLICANT: Lodes, Michael J.
4  APPLICANT: Houghton, Raymond L.
5  APPLICANT: Smith, John M.
6  APPLICANT: McNeill, Patricia D.
7  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION
8  TITLE OF INVENTION: OF T. CRUZI INFECTION
9  FILE REFERENCE: 210121.422C3
10 CURRENT APPLICATION NUMBER: US/09/256,976
11 CURRENT FILING DATE: 1999-02-24
12 NUMBER OF SEQ ID NOS: 95
13 SOFTWARE: PatentIn Ver. 2.0
14 SEQ ID NO 50
15 LENGTH: 207
16 TYPE: PRT
17 ORGANISM: Trypanosoma cruzi
18 US-09-256-976-50

```

Query Match	15.2%	Score 179;	DB 4;	Length 207;
Best Local Similarity	41.0%;	Pred. No. 4.1e-08;		
Matches 43;	Conservative 11;	Mismatches 45;	Indels 6;	Gaps 2

QY 11. ESEFFSAALAPAPPIAAEPAAAAALTTTPVAABPAACAFAAEEPAABAPVAGPAABAP 173
 Db 5 PAAPKSAEKPEAPK-SAPKPAEKPSAKAPKPAEKSGAPKPAEPKSAEPKPAEPKSAEP 63
 QY 174 VAAPPAEAPVGEPAABESP-----ABPATAPAEPEHPSPS 213
 Db 64 KPAPKSAEPKPAEPKSAEPKPAEKSAEPKPAEKSAEPKPAES 108

RESULT 10
US-08-818-112-142
; Sequence 142, Application US/08818112

```

1  GENERAL INFORMATION:
2  APPLICANT: Reed, Steven G.
3  APPLICANT: Skelky, Yaeli A.W.
4  APPLICANT: Dillon, Davin C.
5  APPLICANT: Campos-Neto, Antonio
6  APPLICANT: Houghton, Raymond
7  APPLICANT: Veddick, Thomas S.
8  APPLICANT: Twardzik, Daniel R.
9  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
10 TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
11 NUMBER OF SEQUENCES: 153

```

```

;
;
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS

```


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Avenue, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
FILING DATE: 07-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 142:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:29:52 ; Search time 33 Seconds
(without alignments)
1357.012 Million cell updates/sec

Title: US-09-923-236-2

Perfect score: 219
Sequence: 1 MLLLMACIVCAFAFKRRF.....TAKPAPRPHSPSLEQANQ 219

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 762491 seqs, 204481190 residues

Word size : 0

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	219	100.0	219	9	US-09-922-480-2	Sequence 2, Appl1
2	219	100.0	219	9	US-09-923-236-2	Sequence 2, Appl1
3	219	100.0	219	10	US-09-923-469-2	Sequence 2, Appl1
4	150	68.5	221	12	US-10-231-417-192	Sequence 192, Appl1
5	140	63.9	219	11	US-09-746-783-106	Sequence 106, Appl1
6	9	4.1	798	15	US-10-156-761-13162	Sequence 13162, Appl1
7	9	4.1	2969	10	US-09-738-626-4434	Sequence 4434, Appl1
8	9	3.7	132	12	US-10-289-762-96	Sequence 96, Appl1
9	8	3.7	206	15	US-10-219-220-250	Sequence 250, Appl1
10	8	3.7	249	15	US-10-180-375-82	Sequence 82, Appl1
11	8	3.7	342	12	US-10-369-493-12383	Sequence 12383, Appl1
12	8	3.7	367	12	US-10-369-493-73117	Sequence 73117, Appl1
13	8	3.7	383	10	US-09-788-345-10	Sequence 10, Appl1
14	8	3.7	383	12	US-10-337-312-10	Sequence 10, Appl1
15	8	3.7	412	10	US-09-788-345-12	Sequence 12, Appl1

16	8	3.7	412	12	US-10-337-312-12	Sequence 12, Appl
17	8	3.7	428	10	US-09-906-514-4	Sequence 4, Appl
18	8	3.7	428	12	US-10-177-478-14	Sequence 14, Appl
19	8	3.7	454	15	US-10-156-761-13939	Sequence 13939, A
20	8	3.7	511	9	US-09-864-761-34590	Sequence 34590, A
21	8	3.7	695	12	US-10-029-386-32280	Sequence 32280, A
22	8	3.7	695	12	US-10-369-493-15364	Sequence 15364, A
23	8	3.7	850	12	US-10-369-493-10111	Sequence 10111, A
24	8	3.7	4307	12	US-10-369-493-8698	Sequence 5698, Ap
25	8	3.7	4307	12	US-10-369-493-5699	Sequence 5699, Ap
26	8	3.7	4307	12	US-10-369-493-5700	Sequence 5700, Ap
27	8	3.7	8026	12	US-10-132-151-13	Sequence 12, Appl
28	8	3.7	16	12	US-10-132-151-13	Sequence 43, Appl
29	7	3.2	21	14	US-10-124-800-15	Sequence 15, Appl
30	7	3.2	27	11	US-09-974-879-306	Sequence 306, App
31	7	3.2	27	11	US-09-305-736-306	Sequence 306, App
32	7	3.2	27	12	US-09-818-663-306	Sequence 306, App
33	7	3.2	48	9	US-09-739-907-177	Sequence 177, App
34	7	3.2	48	12	US-09-938-671-177	Sequence 177, App
35	7	3.2	60	9	US-09-864-761-33554	Sequence 33554, A
36	7	3.2	60	9	US-09-864-761-33624	Sequence 33624, A
37	7	3.2	60	9	US-09-864-761-34127	Sequence 34127, A
38	7	3.2	61	9	US-09-864-761-35187	Sequence 35187, A
39	7	3.2	65	9	US-09-864-761-36194	Sequence 36194, A
40	7	3.2	78	11	US-09-820-843A-89	Sequence 89, Appl
41	7	3.2	88	9	US-09-764-853-503	Sequence 503, App
42	7	3.2	91	9	US-09-867-550-738	Sequence 738, App
43	7	3.2	95	9	US-09-739-907-178	Sequence 178, App
44	7	3.2	95	12	US-09-938-671-178	Sequence 178, App
45	7	3.2	105	12	US-10-104-047-2874	Sequence 2874, Ap

ALIGNMENTS

RESULT 1
US-09-922-480-2
Sequence 2, Application US/09922480
Patent No. US20020081701A1
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
TITLE OF INVENTION: SECRETED SALIVARY ZS163 POLYPEPTIDE
FILE REFERENCE: 97-71
CURRENT APPLICATION NUMBER: US/09/922,480
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US 60/124,820
PRIOR FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 219
TYPE: PRT
ORGANISM: Homo sapiens
US-09-922-480-2

QY	Query Match	100.0%	Score 219;	DB 9;	Length 219;
DB	Best Local Similarity	100.0%	Pred. No. 5.6e-185;		
QY	Matches 219;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	MLLLMACIVCAFAFKRRFPPIGDDNDGRLHPSINIPGINLPPPLYRRVNTVP	60		
DB	1	MLLLMACIVCAFAFKRRFPPIGDDNDGRLHPSINIPGINLPPPLYRRVNTVP	60		
QY	61	SYPGNTYDTGSPSPWILTSRGPYVYHNGFPLATOLNVPPLPPRGPFPPSPRFFSA	120		
DB	61	SYPGNTYDTGSPSPWILTSRGPYVYHNGFPLATOLNVPPLPPRGPFPPSPRFFSA	120		
QY	121	AAAPAPPIAAEPAAAPLTATPVAAEPAGAPVAAEPAAEPVGAEPAAEPVAAEPAA	180		
DB	121	AAAPAPPIAAEPAAAPLTATPVAAEPAGAPVAAEPAAEPVGAEPAAEPVAAEPAA	180		
QY	161	EAPVGEPAEPAAEPSPAPRATAKPAPRPHSPSLEQANQ	219		

Db 181 EAPVGEPAAEPSPAEPATAKPAAPDEPHPSLEQANQ 219

RESULT 2
US-09-923-236-2
; Sequence 2, Application US/09923236
; Patent No. US20020090677A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Adler, David A.
; TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
; FILE REFERENCE: 97-71
; CURRENT APPLICATION NUMBER: US/09/923,236
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/124,820
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-236-2

Query Match 100.0%; Score 219; DB 9; Length 219;
Best Local Similarity 100.0%; Pred. No. 5,6e-185;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLMACIVCAFAFKRRFPFIGEDDNDGHLHPSLNIPYGINLPPPLYRPNVTVP 60
DB 1 MLLLMACIVCAFAFKRRFPFIGEDDNDGHLHPSLNIPYGINLPPPLYRPNVTVP 60
QY 61 SYPGNTYTDGTPSYWILTSFGFPYVYHIRGFPPLATOLNVPPLPRGFPVPSRFFSA 120
DB 61 SYPGNTYTDGTPSYWILTSFGFPYVYHIRGFPPLATOLNVPPLPRGFPVPSRFFSA 120
QY 121 AAAPAAPPIAAEPAAAPLTAIPVAEPAAAGAVAAEPAAEPVGAEPAAEPVAAEPAA 180
DB 121 AAAPAAPPIAAEPAAAPLTAIPVAEPAAAGAVAAEPAAEPVGAEPAAEPVAAEPAA 180
QY 181 EAPVGEPAAEPSPAEPATAKPAAPDEPHPSLEQANQ 219
DB 181 EAPVGEPAAEPSPAEPATAKPAAPDEPHPSLEQANQ 219

RESULT 3
US-09-922-469-2
; Sequence 2, Application US/09922469
; Patent No. US20020173027A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Adler, David A.
; TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
; FILE REFERENCE: 97-71
; CURRENT APPLICATION NUMBER: US/09/922,469
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/124,820
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-469-2

Query Match 100.0%; Score 219; DB 10; Length 219;
Best Local Similarity 100.0%; Pred. No. 5,6e-185;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLMACIVCAFAFKRRFPFIGEDDNDGHLHPSLNIPYGINLPPPLYRPNVTVP 60

Db 1 MLLLMACIVCAFAFKRRFPFIGEDDNDGHLHPSLNIPYGINLPPPLYRPNVTVP 60
QY 61 SYPGNTYTDGTPSYWILTSFGFPYVYHIRGFPPLATOLNVPPLPRGFPVPSRFFSA 120
DB 61 SYPGNTYTDGTPSYWILTSFGFPYVYHIRGFPPLATOLNVPPLPRGFPVPSRFFSA 120
QY 121 AAAPAAPPIAAEPAAAPLTAIPVAEPAAAGAVAAEPAAEPVGAEPAAEPVAAEPAA 180
DB 121 AAAPAAPPIAAEPAAAPLTAIPVAEPAAAGAVAAEPAAEPVGAEPAAEPVAAEPAA 180
QY 181 EAPVGEPAAEPSPAEPATAKPAAPDEPHPSLEQANQ 219
DB 181 EAPVGEPAAEPSPAEPATAKPAAPDEPHPSLEQANQ 219

RESULT 4
US-10-231-417-192
; Sequence 192, Application US/10231417
; Publication No. US20030176681A1
; GENERAL INFORMATION:
; APPLICANT: Feng et al.
; TITLE OF INVENTION: 146 Human Secreted Proteins
; FILE REFERENCE: P2019P1
; CURRENT APPLICATION NUMBER: US/10/231,417
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US/09/296,622
; NUMBER OF SEQ ID NOS: 619
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 192
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens

FEATURE: NAME/KEY: SITE
LOCATION: (159)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE: NAME/KEY: SITE
LOCATION: (221)
OTHER INFORMATION: Xaa equals stop translation
US-10-231-417-192

Query Match 68.5%; Score 150; DB 12; Length 221;
Best Local Similarity 100.0%; Pred. No. 3,3e-124;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLMACIVCAFAFKRRFPFIGEDDNDGHLHPSLNIPYGINLPPPLYRPNVTVP 60
DB 1 MLLLMACIVCAFAFKRRFPFIGEDDNDGHLHPSLNIPYGINLPPPLYRPNVTVP 60
QY 61 SYPGNTYTDGTPSYWILTSFGFPYVYHIRGFPPLATOLNVPPLPRGFPVPSRFFSA 120
DB 61 SYPGNTYTDGTPSYWILTSFGFPYVYHIRGFPPLATOLNVPPLPRGFPVPSRFFSA 120
QY 121 AAAPAAPPIAAEPAAAPLTAIPVAEPAA 150
DB 121 AAAPAAPPIAAEPAAAPLTAIPVAEPAA 150

RESULT 5
US-09-746-783-106
; Sequence 106, Application US/09746783
; Publication No. US20030044935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John M.
; Lavallie, Edward R.
; Racine, Lisa A.
; Treacy, Maurice
; Spaulding, Vikki
; Agostino, Michael J.
; Howes, Steven H.

```

; Fecthel, Kim
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridge Park Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/746,783
; FILING DATE: 21-Dec-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Milastincic, Debra J.
; REGISTRATION NUMBER: 46,931
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 219 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 106:
US-09-746-783-106

Query Match          63.9% Score 140; DB 11; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.1e-115;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLIMACTVCVAFARKRFPPFIGEDDNDGPHLPSLNIPIYGINLPPPLYRREVTVP 60
   |||||||
DB 1 MGLIMACTVCVAFARKRFPPFIGEDDNDGPHLPSLNIPIYGINLPPPLYRREVTVP 60

QY 61 SYPGNTYDTGCLPSYPMWITSGSPYVTHIRGEPPLATQNLVPLPPRGEPFVPSRFPESA 120
   |||||||
DB 61 SYPGNTYDTGCLPSYPMWITSGSPYVTHIRGEPPLATQNLVPLPPRGEPFVPSRFPESA 120

QY 121 AAAPAAPPAAPAAAPLT 140
   |||||||
DB 121 AAAPAAPPAAPAAAPLT 140

RESULT 6
US-10-156-761-13162
; Sequence 13162, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02

```

```

; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13162
; LENGTH: 798
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13162

Query Match          4.1% Score 9; DB 15; Length 798;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 PATAKPAAP 206
   |||||||
DB 29 PATAKPAAP 37

RESULT 7
US-09-738-626-4434
; Sequence 4434, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: YATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4434
; LENGTH: 2969
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4434

Query Match          4.1% Score 9; DB 10; Length 2969;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 PAABAPVAA 176
   |||||||
DB 1663 PAABAPVAA 1671

RESULT 8
US-10-289-762-96
; Sequence 96, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Giffels, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 96
; LENGTH: 1332

```

TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-10-289-762-96

Query Match 3.7%; Score 8; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 EAPVAAEP 178
DB 53 EAPVAAEP 60

RESULT 9
US-10-219-220-250
Sequence 250, Application US/10219220
Publication No. US20030082724A1
GENERAL INFORMATION:

APPLICANT: Plim, Barry
APPLICANT: Iasham, Annette
TITLE OF INVENTION: Compositions affecting programmed cell
FILE REFERENCES: 11000.1022c1
CURRENT APPLICATION NUMBER: US/10/219,220
CURRENT FILING DATE: 2002-08-14

PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932
PRIOR FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 290
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 250
LENGTH: 206

TYPE: PRT
ORGANISM: Eucaalyptus grandis
US-10-219-220-250

Query Match 3.7%; Score 8; DB 15; Length 206;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAP 127
DB 112 AAAAPAP 119

RESULT 10
US-10-180-375-82
Sequence 82, Application US/10180375
Publication No. US20030126638A1
GENERAL INFORMATION:

APPLICANT: Allen, William B.
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Famodu, Omolayo O.
APPLICANT: Harwell, Leslie T.
APPLICANT: Helentjaris, Timothy
APPLICANT: Li, Changjiang
APPLICANT: Lowe, Keith
APPLICANT: Oliveira, Igor Cunha
APPLICANT: Shen, Bo

APPLICANT: Tarczyński, Mitchell C.
TITLE OF INVENTION: Alteration of Oil Traits in Plants
FILE REFERENCE: BBI458 US NAI
CURRENT APPLICATION NUMBER: US/10/180,375
CURRENT FILING DATE: 2002-06-26

NUMBER OF SEQ ID NOS: 222
SOFTWARE: Microsoft Office 97
SEQ ID NO 82

LENGTH: 249
TYPE: PRT

ORGANISM: Trilicium aestivum
US-10-180-375-82

Query Match 3.7%; Score 8; DB 15; Length 249;
Best Local Similarity 100.0%; Pred. No. 42;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 VEPAAEP 193
DB 92 VEPAAEP 99

RESULT 11
US-10-369-493-12383
Sequence 12383, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 12383
LENGTH: 342

TYPE: PRT
ORGANISM: Aspergillus nidulans
FEATURE:

NAME/KEY: unsure
LOCATION: (1)-(342)
OTHER INFORMATION: unsure at all Xaa locations

US-10-369-493-12383

Query Match 3.7%; Score 8; DB 12; Length 342;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 SAAAPAA 126
DB 297 SAAAPAA 304

RESULT 12
US-10-369-493-7317
Sequence 7317, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 7317
LENGTH: 367

TYPE: PRT
ORGANISM: Burkholderia cepacia
US-10-369-493-7317

Query Match 3.7%; Score 8; DB 12; Length 367;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 AAGAPVAA 156

Db 71 AAGAPVAA 78

```

RESULT 13
US-09-788-345-10
; Sequence 10, Application US/09788345
; Patent No. US20020147321A1
; GENERAL INFORMATION:
; APPLICANT: ALONSO BEDATE, Carlos
; APPLICANT: REQUEENA ROLANIA, Jose M.
; TITLE OF INVENTION: CHIMERIC GENE FORMED OF THE DNA SEQUENCES THAT ENCODE THE ANTIGEN
; TITLE OF INVENTION: DETERMINANTS OF FOUR PROTEINS OF L. INFANTUM, USEFUL FOR SEROLOG
; TITLE OF INVENTION: CANINE LEISHMANIOSIS AND PROTEIN OBTAINED
; FILE REFERENCE: BEDATE-1A
; CURRENT APPLICATION NUMBER: US/09/788,345
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 09/219,306
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric
US-09-788-345-10

```

```

Query Match 3.7%; Score 8; DB 10; Length 383;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 120 AAAPAP 127
Db 354 AAAPAP 361

```

```

RESULT 14
US-10-337-312-10
; Sequence 10, Application US/10337312
; Publication No. US20030138451A1
; GENERAL INFORMATION:
; APPLICANT: ALONSO BEDATE, Carlos
; APPLICANT: REQUEENA ROLANIA, Jose M.
; TITLE OF INVENTION: CHIMERIC GENE FORMED OF THE DNA SEQUENCES THAT ENCODE THE ANTIGEN
; TITLE OF INVENTION: DETERMINANTS OF FOUR PROTEINS OF L. INFANTUM, USEFUL FOR SEROLOG
; TITLE OF INVENTION: CANINE LEISHMANIOSIS AND PROTEIN OBTAINED
; FILE REFERENCE: BEDATE-1A
; CURRENT APPLICATION NUMBER: US/10/337,312
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US/09/788,345
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 09/219,306
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric
US-10-337-312-10

```

```

Query Match 3.7%; Score 8; DB 12; Length 383;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 120 AAAPAP 127

```

Db 354 AAAPAP 361

```

RESULT 15
US-09-788-345-12
; Sequence 12, Application US/09788345
; Patent No. US20020147321A1
; GENERAL INFORMATION:
; APPLICANT: ALONSO BEDATE, Carlos
; APPLICANT: REQUEENA ROLANIA, Jose M.
; TITLE OF INVENTION: CHIMERIC GENE FORMED OF THE DNA SEQUENCES THAT ENCODE THE ANTIGEN
; TITLE OF INVENTION: DETERMINANTS OF FOUR PROTEINS OF L. INFANTUM, USEFUL FOR SEROLOG
; TITLE OF INVENTION: CANINE LEISHMANIOSIS AND PROTEIN OBTAINED
; FILE REFERENCE: BEDATE-1A
; CURRENT APPLICATION NUMBER: US/09/788,345
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 09/219,306
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: recombinant sequence from vectors pQ31 and pMal
US-09-788-345-12

```

```

Query Match 3.7%; Score 8; DB 10; Length 412;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 120 AAAPAP 127
Db 383 AAAPAP 390

```

Search completed: January 15, 2004, 15:35:27
Job time : 33 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 15:28:37 ; Search time 21 Seconds
(without alignments)
441.242 Million cell updates/sec

Title: US-09-923-236-2

Perfect score: 219
Sequence: 1 MKLLIMACTIVAFARKRRF.....TAKPAAPHPSPSLEQANO 219

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	219	100.0	219	4	US-09-527-345-2
2	9	4.1	322	2	US-08-428-414A-2
3	8	3.7	132	4	US-09-198-452A-96
4	8	3.7	168	4	US-09-252-991A-18199
5	8	3.7	190	4	US-09-252-991A-17963
6	8	3.7	307	4	US-09-252-991A-21588
7	8	3.7	383	4	US-09-471-396-3
8	8	3.7	383	4	US-09-788-345-10
9	8	3.7	410	4	US-09-252-991A-30606
10	8	3.7	412	4	US-09-471-396-1
11	8	3.7	412	4	US-09-788-345-12
12	8	3.7	549	4	US-09-252-991A-20196
13	8	3.7	575	4	US-08-653-648A-5
14	8	3.7	778	4	US-09-252-991A-17591
15	8	3.7	4545	2	US-08-804-227C-14
16	8	3.7	4550	2	US-08-804-227C-8
17	8	3.7	4550	2	US-08-804-198-2
18	8	3.2	31	3	US-09-248-588-27
19	7	3.2	76	6	5273901-11
20	7	3.2	76	6	5482709-10
21	7	3.2	76	6	5486550-6
22	7	3.2	79	4	US-09-252-991A-25964
23	7	3.2	118	4	US-08-301-162-10
24	7	3.2	118	4	US-09-461-240-10
25	7	3.2	118	4	US-09-968-927-10
26	7	3.2	122	4	US-09-328-352-7994
27	7	3.2	134	4	US-09-252-991A-18886

28	7	3.2	136	4	US-09-733-210-814	Sequence 814, App
29	7	3.2	136	4	US-09-733-210-815	Sequence 815, App
30	7	3.2	141	4	US-09-252-991A-26122	Sequence 26122, A
31	7	3.2	156	1	US-08-074-121-5	Sequence 5, Appl1
32	7	3.2	156	5	PCT-US94-06447-5	Sequence 5, Appl1
33	7	3.2	159	3	US-08-481-435-13	Sequence 13, Appl1
34	7	3.2	170	4	US-09-252-991A-17373	Sequence 27069, A
35	7	3.2	173	4	US-09-252-991A-17373	Sequence 17373, A
36	7	3.2	173	4	US-09-252-991A-25321	Sequence 25321, A
37	7	3.2	180	6	5273901-7	Patent No. 5273901
38	7	3.2	180	6	5482709-6	Patent No. 5482709
39	7	3.2	188	4	US-09-252-991A-20399	Sequence 20399, A
40	7	3.2	195	4	US-09-252-991A-20967	Sequence 20967, A
41	7	3.2	197	4	US-09-252-991A-24218	Sequence 24218, A
42	7	3.2	202	4	US-09-252-991A-28326	Sequence 28326, A
43	7	3.2	216	4	US-09-252-991A-18761	Sequence 18761, A
44	7	3.2	239	4	US-09-252-991A-21250	Sequence 21250, A
45	7	3.2	249	4	US-09-252-991A-18733	Sequence 18733, A

ALIGNMENTS

```
RESULT 1
US-09-527-345-2
; Sequence 2, Application US/09527345
; Patent No. 631413
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: SECRETED SALIVARY ZS1G63 POLYPEPTIDE
; FILE REFERENCE: 97-71
; CURRENT FILING DATE: 1999-03-17
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-527-345-2

Query Match      100.0%; Score 219; DB 4; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.1e-203;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKLLIMACTIVAFARKRRFPFIGEDDDGHPILPSINIPYGINLPPPLYRRVNTVP 60
      |||
DB      1 MKLLIMACTIVAFARKRRFPFIGEDDDGHPILPSINIPYGINLPPPLYRRVNTVP 60
      |||

QY      61 SYPGNTYDTGSPSPWILTSPPGPPYVHNGFPLATOLNVPPLPPRGFPVPPSPFSA 120
      |||
DB      61 SYPGNTYDTGSPSPWILTSPPGPPYVHNGFPLATOLNVPPLPPRGFPVPPSPFSA 120
      |||

QY      121 AAPAPAPPIAEPAAAPLTPVAABEPAGAPVAABEPAAEPVABEPAAEPVAAEPAA 180
      |||
DB      121 AAPAPAPPIAEPAAAPLTPVAABEPAGAPVAABEPAAEPVABEPAAEPVAAEPAA 180
      |||

QY      161 EAPVGEPAABEPSPAEPAATKAPAPBPHSPSLEQANO 219
      |||
DB      161 EAPVGEPAABEPSPAEPAATKAPAPBPHSPSLEQANO 219
      |||

RESULT 2
US-08-428-414A-2
; Sequence 2, Application US/08428414A
; Patent No. 5912166
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; LEISHMANIASIS
```

NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,414A
FILING DATE: 21-Apr-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Kadlecsek, Ann T.
REGISTRATION NUMBER: 39,244
REFERENCE/DOCKET NUMBER: 210121.407
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEEDANDBERRY
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 322 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-428-414A-2

Query Match 4.1%; Score 9; DB 2; Length 322;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 AAEPAAAP 138
DB 289 AAEPAAAP 297

RESULT 3
US-09-198-452A-96
Sequence 96, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Grifais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ. ID NOS: 6849
SEQ. ID NO 96
LENGTH: 132
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-96

Query Match 3.7%; Score 8; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 EAPVAAEP 178
DB 53 EAPVAAEP 60

RESULT 4
US-09-252-991A-18199
Sequence 18199, Application US/09252991A

Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ. ID NOS: 33142
SEQ. ID NO 18199
LENGTH: 168
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18199

Query Match 3.7%; Score 8; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 VAAEPAG 151
DB 108 VAAEPAG 115

RESULT 5
US-09-252-991A-17963
Sequence 17963, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ. ID NOS: 33142
SEQ. ID NO 17963
LENGTH: 190
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17963

Query Match 3.7%; Score 8; DB 4; Length 190;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 AAAPAAP 128
DB 113 AAAPAAP 120

RESULT 6
US-09-252-991A-21588
Sequence 21588, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21588
LENGTH: 307
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21588

Query Match 3.7%; Score 8; DB 4; Length 307;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 AAAPAP 128
Db 120 AAAPAP 127

RESULT 7
US-09-471-396-3
Sequence 3, Application US/09471396
Patent No. 6458359
GENERAL INFORMATION:
APPLICANT: BEDATE, Carlos Alonso
APPLICANT: REQUEÑA ROLANTA, Jose Maria
APPLICANT: SOTO ALVAREZ, Manuel
TITLE OF INVENTION: CHIMERIC GENE FORMED OF THE DNA SEQUENCES THAT ENCODE
TITLE OF INVENTION: THE ANTIGENIC DETERMINANTS OF FOUR PROTEINS OF
FILE REFERENCE: bedate2a.seq
CURRENT APPLICATION NUMBER: US/09/471,396
CURRENT FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: 60/113,825
PRIOR FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 3
LENGTH: 383
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Protein Q
US-09-471-396-3

Query Match 3.7%; Score 8; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAPAP 127
Db 354 AAAPAP 361

RESULT 8
US-09-788-345-10
Sequence 10, Application US/09788345
Patent No. 6525186
GENERAL INFORMATION:
APPLICANT: ALONSO BEDATE, Carlos
APPLICANT: REQUEÑA ROLANTA, Jose M.
APPLICANT: SOTO ALVAREZ, Manuel
TITLE OF INVENTION: CHIMERIC GENE FORMED OF THE DNA SEQUENCES THAT ENCODE THE ANTIGEN
TITLE OF INVENTION: DETERMINANTS OF FOUR PROTEINS OF L. INFANTUM, USEFUL FOR SEROLOG
FILE REFERENCE: BEDATE=1A
CURRENT APPLICATION NUMBER: US/09/788,345
CURRENT FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 09/219,306
PRIOR FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patent In version 3.1
SEQ ID NO 10
LENGTH: 383
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chimeric
US-09-788-345-10

Query Match 3.7%; Score 8; DB 4; Length 383;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAPAP 127
Db 354 AAAPAP 361

RESULT 9
US-09-252-991A-30606
Sequence 30606, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubinfeld et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30606
LENGTH: 410
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30606

Query Match 3.7%; Score 8; DB 4; Length 410;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 AAGAPVAA 156
Db 37 AAGAPVAA 44

RESULT 10
US-09-471-396-1
Sequence 1, Application US/09471396
Patent No. 6458359
GENERAL INFORMATION:
APPLICANT: BEDATE, Carlos Alonso
APPLICANT: REQUEÑA ROLANTA, Jose Maria
APPLICANT: SOTO ALVAREZ, Manuel
TITLE OF INVENTION: CHIMERIC GENE FORMED OF THE DNA SEQUENCES THAT ENCODE
TITLE OF INVENTION: THE ANTIGENIC DETERMINANTS OF FOUR PROTEINS OF
FILE REFERENCE: bedate2a.seq
CURRENT APPLICATION NUMBER: US/09/471,396
CURRENT FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: 60/113,825
PRIOR FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1
LENGTH: 412
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Protein Q
US-09-471-396-1

Query Match 3.7%; Score 8; DB 4; Length 412;
Best Local Similarity 100.0%; Pred. No. 8.5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAP 127
|||||

DB 383 AAAAPAP 390

RESULT 11

US-09-788-345-12

Sequence 12, Application US/09788345

Patent No. 6525186

GENERAL INFORMATION:

APPLICANT: ALONSO BEDATE, Carlos

APPLICANT: ROQUELA ROLANTA, Jose M.

APPLICANT: SOTO ALVAREZ, Manuel

TITLE OF INVENTION: CHIMERIC GENE FORMED OF THE DNA SEQUENCES THAT ENCODE THE ANTIGEN

TITLE OF INVENTION: DETERMINANTS OF FOUR PROTEINS OF L. INFANTUM, USEFUL FOR SEROLOG

FILE REFERENCE: BEDATE-1A

CURRENT APPLICATION NUMBER: US/09/788,345

CURRENT FILING DATE: 2001-02-21

PRIOR APPLICATION NUMBER: 09/219,306

PRIOR FILING DATE: 1998-12-23

NUMBER OF SEQ ID NOS: 12

SOFTWARE: Patentin version 3.1

SEQ ID NO 12

LENGTH: 412

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: recombinant sequence from vectors pQ31 and pMal

US-09-788-345-12

Query Match
Best Local Similarity 100.0%; Pred. No. 8.5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 AAAAPAP 127
|||||

DB 383 AAAAPAP 390

RESULT 12

US-09-252-991A-20196

Sequence 20196, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfeld et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 20196

LENGTH: 549

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-20196

Query Match
Best Local Similarity 100.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 AAEPAGA 152
|||||

DB 326 AAEPAGA 333

RESULT 13

US-08-653-648A-5

Sequence 5, Application US/08653648A

Patent No. 6379945

GENERAL INFORMATION:

APPLICANT: Jepson, Ian

APPLICANT: Martine, Andrew

APPLICANT: Martinez, Alberto

TITLE OF INVENTION: A Gene Switch

FILE REFERENCE: PP050047/US

CURRENT APPLICATION NUMBER: US/08/653,648A

CURRENT FILING DATE: 1996-05-24

PRIOR APPLICATION NUMBER: GB 9510759.5

PRIOR FILING DATE: 1995-05-26

PRIOR APPLICATION NUMBER: GB 9605656.9

PRIOR FILING DATE: 1996-03-18

PRIOR APPLICATION NUMBER: GB 9513882.2

PRIOR FILING DATE: 1995-07-07

PRIOR APPLICATION NUMBER: GB 9517316.7

PRIOR FILING DATE: 1995-08-24

NUMBER OF SEQ ID NOS: 65

SOFTWARE: Patentin version 3.0

SEQ ID NO 5

LENGTH: 575

TYPE: PRT

ORGANISM: Helicobacter pylori

US-08-653-648A-5

Query Match
Best Local Similarity 100.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 TATPVAAE 147
|||||

DB 552 TATPVAAE 559

RESULT 14

US-09-252-991A-27591

Sequence 27591, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfeld et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 27591

LENGTH: 778

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-27591

Query Match
Best Local Similarity 100.0%; Pred. No. 15;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 PAAGPVA 155
|||||

DB 655 PAAGPVA 662

RESULT 15

US-08-804-227C-14

Sequence 14, Application US/08804227C

Patent No. 5876991

GENERAL INFORMATION:

APPLICANT: DeHoff, Bradley S.
APPLICANT: Kutschos, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 4545 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-804-227C-14

Query Match 3.7%; Score 8; DB 2; Length 4545;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 AEPAAAP 138
DB 2620 AEPAAAP 2627

Search completed: January 15, 2004, 15:34:05
Job time : 21 secs

THIS PAGE BLANK (USPTO)